

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2003, 12:17:07 ; Search time 4718 Seconds  
(without alignments)  
9517.938 Million cell updates/sec

Title: **US-09-891-138A-1**  
Perfect score: 1543  
Sequence: 1 gctcctggcagagtttcttg.....tgccataataaataatata 1543

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:  
2: gb\_hg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sv:  
13: gb\_un:  
14: gb\_vi:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_lm:  
20: em\_om:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:  
29: em\_vi:  
30: em\_htg\_hum:  
31: em\_htg\_inv:  
32: em\_htg\_other:  
33: em\_htg\_mus:  
34: em\_htg\_pln:  
35: em\_htg\_rod:  
36: em\_htg\_mam:  
37: em\_htg\_vrt:  
38: em\_sy:  
39: em\_htgo\_hum:  
40: em\_htgo\_mus:  
41: em\_htgo\_other:

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1543	100.0	1543	10	AX376573	AX376573 Sequence	
2	1533.4	99.4	1598	6	AF295367	AF295367 Mus muscu	
3	645.8	41.9	60298	2	AC116149	AC116149 Mus muscu	
4	633.6	41.1	60298	2	AC116149	AC116149 Mus muscu	
5	592.4	38.4	1325	9	AF247785	AF247785 Homo sapi	
6	592.4	38.4	1380	9	AF348078	AF348078 Homo sapi	
7	592.4	38.4	1449	9	BC030948	BC030948 Homo sapi	
8	592.4	38.4	1542	6	AX342665	AX342665 Sequence	
9	590.2	38.3	90343	9	AC116026	AC116026 Homo sapi	
10	590.2	38.3	132745	9	AC068647	AC068647 Homo sapi	
11	589.2	38.2	1996	6	AR035943	AR035943 Sequence	
12	588.6	38.1	158144	2	AC016455	AC016455 Homo sapi	
13	147.6	9.6	1977	5	AF031897	AF031897 Meleagris	
14	145.2	9.4	69462	2	AC101335	AC101335 Mus muscu	
15	126.6	8.2	1014	6	AX148186	AX148186 Sequence	
16	126.6	8.2	1014	6	AX379468	AX379468 Sequence	
17	126.6	8.2	1014	6	AX384211	AX384211 Sequence	
18	126.6	8.2	1014	9	AB083598	AB083598 Homo sapi	
19	126.6	8.2	1014	9	AF411109	AF411109 Homo sapi	
20	126.6	8.2	1081	6	AX458238	AX458238 Sequence	
21	126.6	8.2	1092	9	AF370886	AF370886 Homo sapi	
22	126.6	8.2	1414	9	AB065877	AB065877 Homo sapi	
23	126.6	8.2	1729	6	AX191332	AX191332 Sequence	
24	126.6	8.2	9905	6	AX379470	AX379470 Sequence	
25	126.6	8.2	67645	9	AL356486	AL356486 Human DNA	
26	126.6	8.2	156555	9	AC026756	AC026756 Homo sapi	
27	125	8.1	1014	6	AX305130	AX305130 Sequence	
28	125	8.1	1014	6	AX464561	AX464561 Sequence	
29	119.2	7.7	1020	6	AX147840	AX147840 Sequence	
30	114	7.4	136930	5	AL590151	AL590151 Zebrafish	
31	113.6	7.4	1163	5	GDP2Y3	X98283 G. domestic	
32	109	7.1	202838	2	AC108794	AC108794 Mus muscu	
33	108.4	7.0	91107	2	AC116265	AC116265 Rattus no	
34	107	6.9	1074	5	AF069555	AF069555 Meleagris	
35	104	6.7	1020	6	AX458308	AX458308 Sequence	
36	104	6.7	1076	6	AX375230	AX375230 Sequence	
37	103.2	6.7	1498	5	GDATPREC	X73268 G. domestic	
38	102.6	6.6	3204	10	RNU22830	U22830 Rattus norv	
39	101.6	6.6	1308	4	BTU34041	U34041 Bos taurus	
40	101.6	6.6	1666	4	BTP2YRECP	X87628 B. taurus mr	
41	100.8	6.5	1312	9	S81950	S81950 P2 purinoc	
42	100.8	6.5	1426	9	HSU42029	U42029 Human P2Y1	
43	100.8	6.5	2424	9	HSU42030	U42030 Human P2Y1	
44	100.8	6.5	16178	9	HS4569	AJ006945 Human P2Y	
45	100.8	6.5	176170	2	AC021662	AC021662 Homo sapi	

## ALIGNMENTS

RESULT 1  
AX376573  
LOCUS AX376573  
DEFINITION Sequence 1 from Patent WO0200719.  
ACCESSION AX376573  
VERSION AX376573.1 GI:19170674  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1  
AUTHORS Lin, D.C., Zhao, J., Chen, J.L. and Cutler, G.  
TITLE Novel receptors  
JOURNAL Patent: WO 0200719-A 1 03-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

FEATURES		Tularik Inc. (US)		Location/Qualifiers	
source		1. 1543		/organism="Mus musculus"	
		/db_xref="taxon:10090"			
CDS		44..997		/note="mouse TGR18 G-protein coupled receptor (GPCR)"	
		/codon_start=1			
		/protein_id="CAD26816.1"			
		/db_xref="GI:19170675"			
		/translation="MAONLSCENWLATEALINKYLSAFVAIEFIFGLLGNVTVFGY			
		LFCMKNWNSNMYLFNLSIDFAFLPILIKSYANDKGTGVDVLCISNRYVLHNL			
		YTSILFTFTSMYRILYLMKYPFRHFLKQKFAILLSLAWMLVLEVLPMPLFINSV			
		PKESGNCIDYASGNENPHNLISLCLLLGLFLPLSVVMGFFYKRMVFLKRRSQQA			
		TALPLDPKRLVLLAVVIFSLFTPYHIMRNLRIASRLDSWPOGCTQKAISKIYTLIR			
		PLAFNSAINPIFVFLMGDHYREMLISKPRQYKSLTSPRT"			
BASE COUNT		438 a 352 c 293 g 460 t			
ORIGIN		Query Match		100.0%; Score 1543; DB 6; Length 1543;	
		Best Local Similarity		100.0%; Pred. No. 0;	
		Matches 1543; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GCTCCTGGCAGAGTTTCTGTCGACACAGAGCCGACAGAGAGTGGCACAGAAATTATC	60		
DB	1	GCTCCTGGCAGAGTTTCTGTCGACACAGAGCCGACAGAGAGTGGCACAGAAATTATC	60		
QY	61	TTGTGAGAAATGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCAATTTTA	120		
DB	61	TTGTGAGAAATGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCAATTTTA	120		
QY	121	TGCAATCGAGTTCATTTTGGAGTCTGCTGGGAATGCTACTGTTGGTTCGGCTACCTCTT	180		
DB	121	TGCAATCGAGTTCATTTTGGAGTCTGCTGGGAATGCTACTGTTGGTTCGGCTACCTCTT	180		
QY	181	CTGCATGAGAACTGGAACAGCAGAAATGCTATCTTTTAACTCTCTGACTT	240		
DB	181	CTGCATGAGAACTGGAACAGCAGAAATGCTATCTTTTAACTCTCTGACTT	240		
QY	241	TGCTTTCTGTGACCCCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA	300		
DB	241	TGCTTTCTGTGACCCCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA	300		
QY	301	TGGAGATGTTCTGTATAGCAACCGATATGCTTACACCAACCTCTACACCGCAT	360		
DB	301	TGGAGATGTTCTGTATAGCAACCGATATGCTTACACCAACCTCTACACCGCAT	360		
QY	361	CTCTTCCCTCATTATGACCGGATATCTGCTCATGAAAGTACCCCTTTCCGAGA	420		
DB	361	CTCTTCCCTCATTATGACCGGATATCTGCTCATGAAAGTACCCCTTTCCGAGA	420		
QY	421	ACACTTTCTACAAAAGAGAAATTTGCCAATTTAATCTGCTGGTCTGCTGGCCCTTAGT	480		
DB	421	ACACTTTCTACAAAAGAGAAATTTGCCAATTTAATCTGCTGGTCTGCTGGCCCTTAGT	480		
QY	481	GACCTTAGAAGTCTACCCATCTCTTTCATCAATCTGCTCCCAAGAGAGGGCAG	540		
DB	481	GACCTTAGAAGTCTACCCATCTCTTTCATCAATCTGCTCCCAAGAGAGGGCAG	540		
QY	541	TAACTCATCGACTATGCAAGTCTGGAACCCCTGAACACATCTCAITACAGCCCTG	600		
DB	541	TAACTCATCGACTATGCAAGTCTGGAACCCCTGAACACATCTCAITACAGCCCTG	600		
QY	601	CTTGACTTTGTTGGGCTTCTAATTCCTCTCTGTGTGATGCTTCTTCTACTACAAGAT	660		
DB	601	CTTGACTTTGTTGGGCTTCTAATTCCTCTCTGTGTGATGCTTCTTCTACTACAAGAT	660		
QY	661	GGTAGTCTTTAAAGAGAGGAGCCAGCAGCAACCTGCCCTGCCACTGGACAAACC	720		
DB	661	GGTAGTCTTTAAAGAGAGGAGCCAGCAGCAACCTGCCCTGCCACTGGACAAACC	720		
QY	721	CCAAACGCTGGTGGTCTGGCGGTTGTGATCTTCTATACCTCTTACACCCCTATCATAT	780		

RESULT 2  
AF295367  
LOCUS AF295367 1598 bp mRNA linear ROD 06-APR-2001  
DEFINITION Mus musculus G-protein coupled receptor GPR91 mRNA, complete cds.  
ACCESSION AF295367  
VERSION AF295367.1 GI:12711490  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 1598)  
AUTHORS Wittenberger, T., Schaller, H.C. and Hellebrand, S.  
TITLE An expressed sequence tag (EST) data mining strategy succeeding in  
the discovery of new G-protein coupled receptors  
JOURNAL J. Mol. Biol. 307 (3), 799-813 (2001)  
MEDLINE 21172992  
PUBMED 11273702

REFERENCE	2 (bases 1 to 1598)
AUTHORS	Wittenberger T., Schaller C.H. and Hellebrand S.
TITLE	Direct Submission
JOURNAL	Submitted (14-AUG-2000) ZMNH, Institut fur Entwicklungsneurobiologie, Martinistr. 52, Hamburg 20246, Germany
FEATURES	Location/Qualifiers
source	1..1598
	/organism="Mus musculus"
	/strain="C57BL"
CDS	/db_xref="taxon:10090"
	74..1027
	/note="orphan receptor"
	/codon_start=1
	/product="G-protein coupled receptor GPR91"
	/protein_id="AAK01867.1"
	/db_xref="GI:12711491"
	/translation="MAQNLSCENLATEALINKYVLSAFVAIEFIFGLGNVTIVRGY LFCKMNWSNVLYLFNLSISDFAPLCTLPILIKSYANDKGTVGDVLICSNRYVHLTNL YTSMLLTYSMBRYLMKYTFREHFLQKREFALLISLAWALVTLEVPMLPTFFINSV PKREGSCIDYASNGPEHNLIYSLCTLGLTLPISVMCFYFKMVVFLKKRSQQQA TLAPDKPQLRVLAIVFISELPYHMRLRIASRLDSWPQGCTOKAKISKIYITLTR PLAFLNSAINPIFYLVFGDHYRMLISKFRPYKSLTSFT"
BASE COUNT	465 a 303 g 472 t
ORIGIN	
Query Match	99.4%; Score 1533.4; DB 10; Length 1598;
Best Local Similarity	99.6%; Pred. No. 0;
Matches 1537; Conservative	0; Mismatches 6; Indels 0; Gaps 0;
Qy	1 GCTCCTGGCAGATTTCTGTGCAGACAGAAGCGGACGAGAAATGCCAGAAATTATC 60
Db	31 GCTCCTGGCAGATTTCTGTGCAGACAGAAGCGGACGAGAAATGCCAGAAATTATC 90
Qy	61 TTGTGAGAAATGGTGTGCAACAGAGGCTATCTTGAATAAGTAAGTACTACCTCTCTGCATTTTA 120
Db	91 TTGTGAGAAATGGTGTGCAACAGAGGCTATCTTGAATAAGTAAGTACTACCTCTCTGCATTTTA 150
Qy	121 TGCAATCGAGTTCATTTTGGACTGCTTGGAAATGCTGTGGTGTTGGCTACCTCTT 180
Db	151 TGCAATCGAGTTCATTTTGGACTGCTTGGAAATGCTGTGGTGTTGGCTACCTCTT 210
Qy	181 CTGCATGAAGAAGCTGGAACAGCAGCAATGCTATCTTTTTAACTTTCCATCTCTGACCTT 240
Db	211 CTGCATGAAGAAGCTGGAACAGCAGCAATGCTATCTTTTTAACTTTCCATCTCTGACCTT 270
Qy	241 TGCTTTCTGTGCACCCTTCCCATTCTGATATAAGAGTATGCCAATGATAAGGGGACCTA 300
Db	271 TGCTTTCTGTGCACCCTTCCCATTCTGATATAAGAGTATGCCAATGATAAGGGGACCTA 330
Qy	301 TGGAGATGTTCTCTGTATAGCAACCGATATGCTTTCACACAACCTCTACACAGCAT 360
Db	331 TGGAGATGTTCTCTGTATAGCAACCGATATGCTTTCACACAACCTCTACACAGCAT 390
Qy	361 CCTTTCTCTCACCTTTCATTAGCATGACCGGATATCTTGCTCATGAAGTACCCCTTCCGAGA 420
Db	391 GCCTTGGCTCACTGTCATTAGCATGACCGGATATCTGCTCATGAAGTACCCCTTCCGAGA 450
Qy	421 ACACCTTCTACAAAAGAAGAAFTTGCCATTTTAATCTCGCTGGCTGTCTGGGCTTACT 480
Db	451 ACACCTTCTACAAAAGAAGAAFTTGCCATTTTAATCTCGCTGGCTGTCTGGGCTTACT 510
Qy	481 GACCTTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCCTGCCCCAAAAAGAGGGCAG 540
Db	511 GACCTTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCCTGCCCCAAAAAGAGGGCAG 570
Qy	541 TAACTGCATCGCATATGCAAGTCTTGGAAACCCCTGAAACAATCTCATTTACAGCCTCTG 600
Db	571 TAACTGCATCGCATATGCAAGTCTTGGAAACCCCTGAAACAATCTCATTTACAGCCTCTG 630
Qy	601 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACCAAGT 660
Db	631 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACCAAGT 690

Thu Apr 3 10:42:25 2003

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP24-540E9

2 (bases 1 to 60298)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boukhalil,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,I.,  
Mihova,T., Mieng,A., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L24912  
Center clone name: 540\_E-9

-----  
NOTE: This record contains 77 individual  
sequencing reads that have not been assembled into  
contigs. Runs of N are used to separate the reads  
and the order in which they appear is completely  
arbitrary. Low-pass sequence sampling is useful for  
identifying clones that may be gene-rich and allows  
overlap relationships among clones to be deduced.  
However, it should not be assumed that this clone  
will be sequenced to completion. In the event that  
the record is updated, the accession number will  
be preserved.

1  
657 756: gap of 100 bp  
757 1426: contig of 670 bp in length  
1427 1526: gap of 100 bp  
1527 2210: contig of 684 bp in length  
2211 2310: gap of 100 bp  
2311 2997: contig of 687 bp in length  
2998 3097: gap of 100 bp  
3098 3786: contig of 689 bp in length  
3787 3886: gap of 100 bp  
3887 4577: contig of 691 bp in length  
4578 4677: gap of 100 bp  
4678 5357: contig of 680 bp in length  
5358 5457: gap of 100 bp  
5458 6150: contig of 693 bp in length  
6151 6250: gap of 100 bp  
6251 6817: contig of 567 bp in length  
6818 6917: gap of 100 bp  
6918 7615: contig of 698 bp in length  
7616 7715: gap of 100 bp  
7716 8412: contig of 697 bp in length

8413 8512: gap of 100 bp  
8513 9198: contig of 686 bp in length  
9199 9298: gap of 100 bp  
9299 9988: contig of 690 bp in length  
9989 10088: gap of 100 bp  
10089 10768: contig of 680 bp in length  
10769 10868: gap of 100 bp  
10869 11524: contig of 656 bp in length  
11525 11624: gap of 100 bp  
11625 12242: contig of 618 bp in length  
12243 12342: gap of 100 bp  
12343 13040: contig of 698 bp in length  
13041 13140: gap of 100 bp  
13141 13829: contig of 689 bp in length  
13830 13929: gap of 100 bp  
13930 14647: contig of 718 bp in length  
14648 14747: gap of 100 bp  
14748 15451: contig of 704 bp in length  
15452 15551: gap of 100 bp  
15552 16247: contig of 696 bp in length  
16248 16347: gap of 100 bp  
16348 17028: contig of 681 bp in length  
17029 17128: gap of 100 bp  
17129 17802: contig of 674 bp in length  
17803 17902: gap of 100 bp  
17903 18593: contig of 691 bp in length  
18594 18693: gap of 100 bp  
18694 19375: contig of 682 bp in length  
19376 19475: gap of 100 bp  
19476 20082: contig of 607 bp in length  
20083 20182: gap of 100 bp  
20183 20875: contig of 693 bp in length  
20876 20975: gap of 100 bp  
20976 21650: contig of 675 bp in length  
21651 21750: gap of 100 bp  
21751 22427: contig of 677 bp in length  
22428 22527: gap of 100 bp  
22528 23238: contig of 711 bp in length  
23239 23338: gap of 100 bp  
23339 24028: contig of 690 bp in length  
24029 24128: gap of 100 bp  
24129 24803: contig of 675 bp in length  
24804 24903: gap of 100 bp  
24904 25603: contig of 700 bp in length  
25604 25703: gap of 100 bp  
25704 26357: contig of 654 bp in length  
26358 26457: gap of 100 bp  
26458 27140: contig of 683 bp in length  
27141 27240: gap of 100 bp  
27241 27946: contig of 706 bp in length  
27947 28046: gap of 100 bp  
28047 28734: contig of 688 bp in length  
28735 28834: gap of 100 bp  
28835 29536: contig of 702 bp in length  
29537 29636: gap of 100 bp  
29637 30324: contig of 688 bp in length  
30325 30424: gap of 100 bp  
30425 31130: contig of 706 bp in length  
31131 31230: gap of 100 bp  
31231 31910: contig of 680 bp in length  
31911 32010: gap of 100 bp  
32011 32691: contig of 681 bp in length  
32692 32791: gap of 100 bp  
32792 33482: contig of 691 bp in length  
33483 33582: gap of 100 bp  
33583 34274: contig of 692 bp in length  
34275 34374: gap of 100 bp  
34375 35081: contig of 707 bp in length  
35082 35181: gap of 100 bp  
35182 35861: contig of 680 bp in length  
35862 35961: gap of 100 bp  
35962 36660: contig of 699 bp in length  
36661 36760: gap of 100 bp





COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L24912  
 Center clone name: 540\_E\_9

\* NOTE: This record contains 77 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1 656: contig of 656 bp in length  
 657 756: gap of 100 bp  
 757 1426: contig of 670 bp in length  
 1427 1526: gap of 100 bp  
 1527 2210: contig of 684 bp in length  
 2211 2310: gap of 100 bp  
 2311 2997: contig of 687 bp in length  
 2998 3097: gap of 100 bp  
 3098 3786: contig of 689 bp in length  
 3787 3886: gap of 100 bp  
 3887 4577: contig of 691 bp in length  
 4578 4677: gap of 100 bp  
 4678 5357: contig of 680 bp in length  
 5358 5457: gap of 100 bp  
 5458 6150: contig of 693 bp in length  
 6151 6250: gap of 100 bp  
 6251 6817: contig of 567 bp in length  
 6818 6917: gap of 100 bp  
 6918 7615: contig of 698 bp in length  
 7616 7715: gap of 100 bp  
 7716 8412: contig of 697 bp in length  
 8413 8512: gap of 100 bp  
 8513 9198: contig of 686 bp in length  
 9199 9298: gap of 100 bp  
 9299 9988: contig of 690 bp in length  
 9989 10088: gap of 100 bp  
 10089 10768: contig of 680 bp in length  
 10769 10868: gap of 100 bp  
 10869 11524: contig of 656 bp in length  
 11525 11624: gap of 100 bp  
 11625 12242: contig of 618 bp in length  
 12243 12342: gap of 100 bp  
 12343 13040: contig of 698 bp in length  
 13041 13140: gap of 100 bp  
 13141 13829: contig of 689 bp in length  
 13830 13929: gap of 100 bp  
 13930 14647: contig of 718 bp in length  
 14648 14747: gap of 100 bp  
 14748 15451: contig of 704 bp in length  
 15452 15551: gap of 100 bp  
 15552 16247: contig of 696 bp in length  
 16248 16347: gap of 100 bp  
 16348 17028: contig of 681 bp in length  
 17029 17128: gap of 100 bp  
 17129 17802: contig of 674 bp in length  
 17803 17902: gap of 100 bp  
 17903 18593: contig of 691 bp in length  
 18594 18693: gap of 100 bp  
 18694 19375: contig of 682 bp in length

19376 19475: gap of 100 bp  
 19476 20082: contig of 607 bp in length  
 20083 20182: gap of 100 bp  
 20183 20875: contig of 693 bp in length  
 20876 20975: gap of 100 bp  
 20976 21650: contig of 675 bp in length  
 21651 21750: gap of 100 bp  
 21751 22427: contig of 677 bp in length  
 22428 22527: gap of 100 bp  
 22528 23238: contig of 711 bp in length  
 23239 23338: gap of 100 bp  
 23339 24028: contig of 690 bp in length  
 24029 24128: gap of 100 bp  
 24129 24803: contig of 675 bp in length  
 24804 24903: gap of 100 bp  
 24904 25603: contig of 700 bp in length  
 25604 25703: gap of 100 bp  
 25704 26357: contig of 654 bp in length  
 26358 26457: gap of 100 bp  
 26458 27140: contig of 683 bp in length  
 27141 27240: gap of 100 bp  
 27241 27946: contig of 706 bp in length  
 27947 28046: gap of 100 bp  
 28047 28734: contig of 688 bp in length  
 28735 28834: gap of 100 bp  
 28835 29536: contig of 702 bp in length  
 29537 29636: gap of 100 bp  
 29637 30324: contig of 688 bp in length  
 30325 30424: gap of 100 bp  
 30425 31130: contig of 706 bp in length  
 31131 31230: gap of 100 bp  
 31231 31910: contig of 680 bp in length  
 31911 32010: gap of 100 bp  
 32011 32691: contig of 681 bp in length  
 32692 32791: gap of 100 bp  
 32792 33482: contig of 691 bp in length  
 33483 33582: gap of 100 bp  
 33583 34274: contig of 692 bp in length  
 34275 34374: gap of 100 bp  
 34375 35081: contig of 707 bp in length  
 35082 35181: gap of 100 bp  
 35182 35861: contig of 680 bp in length  
 35862 35961: gap of 100 bp  
 35962 36660: contig of 699 bp in length  
 36661 36760: gap of 100 bp  
 36761 37447: contig of 687 bp in length  
 37448 37547: gap of 100 bp  
 37548 38243: contig of 696 bp in length  
 38244 38343: gap of 100 bp  
 38344 39034: contig of 691 bp in length  
 39035 39134: gap of 100 bp  
 39135 39813: contig of 679 bp in length  
 39814 39913: gap of 100 bp  
 39914 40597: contig of 684 bp in length  
 40598 40697: gap of 100 bp  
 40698 41392: contig of 695 bp in length  
 41393 41492: gap of 100 bp  
 41493 42190: contig of 698 bp in length  
 42191 42290: gap of 100 bp  
 42291 42967: contig of 677 bp in length  
 42968 43067: gap of 100 bp  
 43068 43736: contig of 669 bp in length  
 43737 43836: gap of 100 bp  
 43837 44525: contig of 689 bp in length  
 44526 44625: gap of 100 bp  
 44626 45306: contig of 681 bp in length  
 45307 45406: gap of 100 bp  
 45407 46111: contig of 705 bp in length  
 46112 46211: gap of 100 bp  
 46212 46848: contig of 637 bp in length  
 46849 46948: gap of 100 bp  
 46949 47639: contig of 691 bp in length  
 47640 47739: gap of 100 bp





```

Db 890 TGGCTTTCTGAAAGAGTGTATCAACCCCTGCTCTCTATTTCTTTGGGAGATCACTTCA 949
Qy 936 GAGAGATGCTGATAGTAAGTTCAGACAATATCTCAAGTCCCTTACATCTCTCAGGACAT 995
Db 950 GGGACATGCTGATGAATCACTAGACACAACCTTCAAACTCCCTTATCATCTTTAGCAGAT 1009
Qy 996 GAGTGTGATGATGAGTCTTCACTCAGCCAAA-TGACACACTTGATAAACAG 1048
Db 1010 GGGCTATGAACCTCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAACAG 1063

RESULT 7
BC030948 1449 bp mRNA linear PRI 13-JUN-2002
LOCUS Homo sapiens, G protein-coupled receptor 91, clone MGC:32514
DEFINITION IMAGE:4594810, mRNA, complete cds.
ACCESSION BC030948
VERSION BC030948.1 GI:21410927
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1449)
Strausberg,R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA.
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 41 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14780893.
FEATURES
Location/Qualifiers
1..1449
/organism="Homo sapiens"
/db_xref="LocusID:56670"
/db_xref="taxon:9606"
/clone="MGC:32514 IMAGE:4594810"
/tissue_type="kidney"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B"
/note="vector: pDNR-LIB"
100..1104
/codon_start=1
/product="G protein-coupled receptor 91"
/protein_id="AAH30948.1"
/db_xref="GI:21410928"
/translation="MLGIMAWNATCKNWLAAEALEKYLISIFVIEFVVGVLGNTIV
VYGFISLKNWNSNIVLENLSVSLAFLCTPLMIBSYANGNWIQDVLCISNRVYL
HANYLTSTLELTFIDIRYLLIKYPFREHLLOKKEFALLSLAIWLVLLELLPLPL
INPVITDGTNDFASSGPPNLIYSMLCTLDGLFLIFVIMCFYFKTALFLKORN
QVATLPLKPLNLIIMAVIFSVLTFYHVRNRIASRLGWSKQYQCTQVINSF
YIVTRPLAFINSVINPVYFLLGDHFRDMLNOLRHNFKSLTFSRWAHELLLSFREK
"

BASE COUNT 411 a 308 c 287 g 443 t
ORIGIN

```

```

Query Match 38.4%; Score 592.4; DB 9; Length 1449;
Best Local Similarity 75.3%; Pred. No. 6.5e-123;
Matches 764; Conservative 0; Mismatches 246; Indels 4; Gaps 2;

Qy 39 GCAGAAATGCGACAGAAATTTATCTTGAGAAATTTGCTTGGCAACAGAGGCTATCTTGAATA 98
Db 107 GGATCATGCGATGGAATGCAACTTGC AAAAAGCTGCTGGCAGCAGAGGCTGCCCTGGAAA 166
Qy 99 AGTACTACTCTCTCGCATTTTATGCAATCGAGTTCAATTTTGGAGCTGCTGGGAATGCA 158
Db 167 AGTACTACTCTTCCATTTTATGGGATTTAGTTCTGTTGGGAGTCTCTTGAATAACCA 226
Qy 159 CTGTGGTGTGGCTTACCTCTCTCATGAAGAAGTGAACAGCAGCAATCTCTATCTTT 218
Db 227 TTGTTGTTTACGGCTACATCTCTCTGAAGAAGTGAACAGCAGCAATATTTATCTCT 286
Qy 219 TTAACTTTTCCATCTCTCACTTTGCTTCTGTGACACCTTCCCATCTCTGATAAGAGTT 278
Db 287 TTAACTCTCTCTCTCACTTAGCTTTCTGTGACACCTCCCATCTCTGATAAGAGTT 346
Qy 279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATGAAGCAACCGATATGCTTC 338
Db 347 ATGCCAATGGAAGAACTGGATATATGAGACGCTCTGTCATGAAGCAACCGATATGCTTC 406
Qy 339 ACACCAACTCTACACAGCATCTCTCTCTCACTTTTCAATAGCATGGACCGATATCTGC 398
Db 407 ATGCCAACTCTATACAGCAATCTCTCTCTCACTTTTATCAGCATAGATGATCTTGA 466
Qy 399 TCATGAAGTACCTTTCCGAGAACACTTTCTACAAAAGAGAAATTTGCCATTTTAATCT 458
Db 467 TAAATAGTATCTTTCCGAGAACACTTTCTGCAAAAAGAAAGATTTGCTATTTTAATCT 526
Qy 459 CGCTGGCTGTCTGGGCTTAGTACCTTAGAAGTTCTACCATCTCTCACTTTTCAATCAATT 518
Db 527 CTTTGGCCATTTGGGCTTTAGTAACCTTAGAGTTACTACCATCTCTTCCCTTATAATC 586
Qy 519 CTGTCCCAAAAAGAGAGGGCAGTAGTAACCTGATGCAAGTTCTGGAAGTTCTGGAACCCCTGAAC 578
Db 587 CTGTATAACTTGACAATGGCACCACCTGTATGATTTTGCAGTTCTGGAGACCCCAACT 646
Qy 579 ACAATCTATTTACAGCCTCTGCTGACTTTGTTGGGCTTCTCAATTTCTCTCTGTA 638
Db 647 ACAACCTCATTTACAGCATGTGCTTAACACTGTTGGGCTTCTTATTTCTCTTTTGTGA 706
Qy 639 TGTCTCTCTCTACTACAAGATGCTACTCTTCTTAAGAGAGAGAGCAGCAGCAACAA 698
Db 707 TGTGTTTCTTTTATACAGATGCTCTCTTCTTAAAGCAGAGAAATAGGAGGTTGCTA 766
Qy 699 CTGCCCTGCCACTGGACAAAACCCCAAGCCCTGCTGTTGGTCTCTCTCTCTCTCTA 758
Db 767 CTGCTCTGCCCTTTGAAAAGCCTCTCAACTTTGGTCACTATGGCAGTGTATCTCTCTG 826
Qy 759 TACTCTTACACCCCTATCATATCATGCGCAATTTGAGATCGGCTCACGCTCGATGTT 818
Db 827 TGTCTTTTACACCCCTATCAGTCTATGCGGAATGTGAGGATCGGCTTCAAGCTGGGAGTT 886
Qy 819 G---GCCACAAGGATGTACACAGAGGCCATCAAACTATATACACTCACACGCGCTC 875
Db 887 GGAAGCAGTATCAGTGCACCTCAGTCTGTCATCAACTCTCTTTTACATTGTACACGGCTT 946
Qy 876 TGGCTTTTCTGAACAGTGGCCTCAATCCCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTA 935
Db 947 TGGCTTTTCTGAACAGTGTCTATCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 1006
Qy 936 GAGAGATGCTGATTAGTAAGTTCAGACAATPACTTCAAGTCCCTTTACATCTCTCAGACAT 995
Db 1007 GGGACATGCTGATGAATCAACTGAGACACAACCTTCAAACTCCCTTACATCTCTTAGCAGAT 1066
Qy 996 GAGTGTGATGAGTCTCTCACTCAGCAAAA-TGAGACACTTTGATAAACAG 1048
Db 1067 GGGCTCATGAACCTCTACTTTTCAATTCAGAGAAAAGTGAAGGGCTTGTGAACAG 1120

```



Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E.,  
 Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,  
 Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,  
 Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L.,  
 Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,  
 Warren, R., Washington, C., Watlington, S., Williams, G.,  
 Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,  
 Wu, Y. F., Zhou, J., Zorrilla, S., Naylor, S. L., Weinstein, G. and  
 Gibbs, R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 90343)  
 Worley, K.C.

Direct Submission  
 Submitted (23-MAR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 90343)  
 Worley, K.C.

Direct Submission  
 Submitted (09-APR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintained sequence  
 continuity across the splice junctions. Sequences that are not  
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
 standard of double strand coverage with a minimum of 2 clones and 2  
 reads with no ambiguities or 2 chemistries with a minimum of 2  
 clones and 3 reads with no ambiguities. If the sequence quality for  
 a region does not meet this standard, it will be indicated in the  
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
 standards - estimated error rate less than 1 per 10,000 bases.  
 Reports of lowest quality individual bases and measures of base  
 quality are listed below. Description of the metrics can be found  
 at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

#### QUALSTAT-REPORT.

Location/Qualifiers  
 1. .90343  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="RP11-3F11"  
 991. .1106  
 /rpt\_family="MER45B"  
 complement(1314. .1627)  
 /rpt\_family="AluSx"  
 complement(2137. .2430)  
 /rpt\_family="AluY"

repeat\_region  
 complement(2568. .2741)  
 /rpt\_family="L1M4"  
 complement(2742. .3047)  
 /rpt\_family="AluSx"  
 complement(3048. .3165)  
 /rpt\_family="L1M4"  
 4735. .4865  
 /rpt\_family="FLAM\_C"  
 5657. .5762  
 /rpt\_family="L1MC/D"  
 5906. .6237  
 /rpt\_family="L1R21B"  
 6289. .6773  
 /rpt\_family="HERVPH21"  
 complement(8725. .9597)  
 /rpt\_family="MER11D"  
 12399. .12689  
 /standard\_name="136046"  
 13774. .13816  
 /rpt\_family="Alu"  
 13817. .13874  
 /rpt\_family="(TA)n"  
 complement(15157. .15633)  
 /rpt\_family="L2"  
 15706. .15747  
 /rpt\_family="AT\_rich"  
 16025. .16235  
 /rpt\_family="MIR"  
 16360. .16682  
 /rpt\_family="L2"  
 complement(16710. .17265)  
 /rpt\_family="L1R49"  
 18077. .18368  
 /rpt\_family="AluSx"  
 complement(18376. .18471)  
 /rpt\_family="L2"  
 complement(18486. .18859)  
 /rpt\_family="MER57B"  
 complement(20618. .20922)  
 /rpt\_family="AluSx"  
 21337. .21363  
 /rpt\_family="AT\_rich"  
 22155. .22561  
 /rpt\_family="L1M4"  
 complement(22608. .22659)  
 /rpt\_family="L1M4"  
 22685. .23013  
 /rpt\_family="L1MB8"  
 23103. .23399  
 /rpt\_family="AluSg"  
 23500. .23973  
 /rpt\_family="L1MB3A"  
 complement(24027. .24305)  
 /rpt\_family="L1MB1"  
 24304. .24655  
 /rpt\_family="L1MB3A"  
 24656. .24697  
 /rpt\_family="MADE1"  
 25203. .25518  
 /rpt\_family="AluJo"  
 25783. .25817  
 /rpt\_family="(TAA)n"  
 26187. .26211  
 /rpt\_family="AT\_rich"  
 27014. .27030  
 /rpt\_family="AT\_rich"  
 complement(27031. .27316)  
 /rpt\_family="AluSx"  
 27317. .27328  
 /rpt\_family="AT\_rich"  
 27574. .27615  
 /rpt\_family="(TAGA)n"  
 28062. .28166  
 STS





REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 132745)  
Worley, K.C.

Direct Submission  
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 132745)  
Worley, K.C.

Direct Submission  
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 132745)  
Worley, K.C.

Direct Submission  
Submitted (29-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

6 (bases 1 to 132745)  
Worley, K.C.

Direct Submission  
Submitted (25-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

7 (bases 1 to 132745)  
Worley, K.C.

Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Mar 28, 2002 this sequence version replaced gi:19718616.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT REPORT.

Location/Qualifiers  
1. .132745  
/organism="Homo sapiens"

FEATURES  
source

```

/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-64D22"
1..2005
/notice="overlaps bases 170209..172213 of clone AC069067"
30..130
/standard_name="74493"
complement(522..1015)
/rpt_family="MLTID"
complement(2452..2697)
/rpt_family="L1MA5A"
complement(3200..3578)
/rpt_family="MLT1B"
3600..3749
/rpt_family="(TA)n"
4391..4411
/rpt_family="AT_rich"
4909..4960
/rpt_family="AT_rich"
complement(5657..6403)
/rpt_family="L1PA13"
6404..7799
/rpt_family="L1PA13"
8045..8318
/standard_name="183647"
complement(8708..9282)
/rpt_family="L1MD3"
complement(9287..9357)
/rpt_family="MLTIF1"
complement(9359..9460)
/rpt_family="L1MC3"
complement(9587..9880)
/rpt_family="AluSg"
complement(10203..10450)
/rpt_family="L1MC4"
11468..11699
/rpt_family="L1M4"
11717..11886
/rpt_family="MLT2CB"
11508..11982
/rpt_family="Tigger3(Golem)"
12020..12246
/rpt_family="THE1C"
12263..12562
/rpt_family="AluSx"
13326..13346
/rpt_family="AT_rich"
complement(13934..14245)
/rpt_family="AluSx"
complement(14256..14568)
/rpt_family="AluJo"
14618..14746
/rpt_family="MER8"
14825..14849
/rpt_family="AT_rich"
14865..14906
/rpt_family="AT_rich"
15465..15739
/rpt_family="L2"
16579..16756
/rpt_family="(TTATA)n"
complement(16757..17074)
/rpt_family="L2"
17621..17660
/rpt_family="(CAAAA)n"
18544..18725
/rpt_family="MIR"

```

## Query Match

38.3%; Score 590.2; DB 9; Length 132745;  
Best Local Similarity 75.5%; Pred. No. 2.3e-122;  
Matches 760; Conservative 0; Mismatches 243; Indels 4; Gaps 2;







RESULT 13	AF031897	1977 bp	mRNA	linear	VRT 06-JAN-1998
LOCUS	Meleagris gallopavo G protein coupled p2Y nucleotide receptor				
DEFINITION	(p2Y) mRNA, complete cds.				
ACCESSION	AF031897				
VERSION	AF031897.1	GI:2707255			
KEYWORDS	Meleagris gallopavo.				
SOURCE	Meleagris gallopavo.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Meleagridae;				
REFERENCE	1 (bases 1 to 1977)				
AUTHORS	Boyer, J.L., Waldo, G.L. and Harden, T.K.				
TITLE	Molecular cloning and expression of an avian G protein-coupled p2Y receptor				
JOURNAL	Mol. Pharmacol. 52 (6), 928-934 (1997)				
MEDLINE	98086419				
PUBMED	9415702				
REFERENCE	2 (bases 1 to 1977)				
AUTHORS	Boyer, J.L., Waldo, G.L. and Harden, T.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-OCT-1997) Dept. of Pharmacology, University of North Carolina School of Medicine, CB# 7365 Mary Ellen Jones Building, Chapel Hill, NC 27599-7365, USA				
FEATURES	Location/Qualifiers				
source	1..1977				
gene	/organism="Meleagris gallopavo"				
CDS	/db_xref="taxon:9103"				
	/tissue_type="blood"				
	1..1977				
	/gene="tp2y"				
	374..1498				
	/gene="tp2y"				
	/codon_start=1				
	/product="G protein coupled p2Y nucleotide receptor"				
	/protein_id="AAC60339.1"				
	/db_xref="GI:2707256"				
	/translation="MDAPVRFSLAPWTPTPTPLWGGNTAAAEAKVNEEFKILL				
	PTISGVFVVGGLPLNSAMWIVSVRPMNATYFNLAISDTLYVSLPLVYYA				
	DRNNPFGVFKIVRFLEYANLYSILELTCISVHRMGICHPILRSUKWIKHARL				
	ICGVWVYTCILPLNLEIETVTSKDNSTLCHDTTKPEFDHYHYSIMALLFQIP				
	FLVGVVYICMAKRKIFSPSPVPSKRSIKMIIIVLTFVAFICFVPEPIHRTL				
	YTSRYFADQCLNINFTYKTRPLASINCLDPILYFMAGDKYRGLRRRAQRP				
	RPVPTSLALVSPVSDSVVGVGSCNSESRGMTVNSRGQ"				
BASE COUNT	416 a 607 c 531 g 423 t				
ORIGIN					
Query Match	9.6%; Score 147.6; DB 5; Length 1977;				
Best Local Similarity	51.9%; Pred. No. 9.7e-23;				
Matches	436; Conservative 0; Mismatches 389; Indels 15; Gaps 4;				
QY	119 TATGCAATCAGTCTATTTTGGACCTGTTGGGAATGTCACCTGGTGTGGCTACCTC	178			
Db					
QY	515 TAGGCATCTCTCTGGTGGTCTGCGCCCTCACTCTCGGCCATGTGGATTTTGTG	574			
Db					
QY	179 TTCTGCATGAAGAACTGGACAGCAGCAATGTCTATCTTTTAACTTTCCATCTCTGAC	238			
Db					
QY	575 TCCAGATGAGGCTTGGGAATGCCACCTACTTACATGTTCAACCTGGCTATCTGAC	634			
Db					
QY	239 TTGTCTTCTGTGACACCTTCCCATCTCTGATAAAGAGTTATGCCAATG---ATAAGGG	295			
Db					
QY	635 ACACCTATCTCTCTCTCCCTCCCAACCCCTGCTACTATTATCTGACCGCAACACTGG	694			
Db					
QY	296 ACCTATGGAGATCTCTCTGTATACCAACCGATATGCTCTCACACCAACCTCTACAC	355			
Db					
QY	695 CCCTCGGAAAGTGTCTGCAAGATCTGCGCTTCTCTACGCCAACCTCTACACG	754			
Db					
QY	356 AGCATCT	415			
Db					
QY	755 AGCATCT	814			
Db					
QY	416 CGAGAACACTTTTACAAAGAGAGGAATTTGCCATTTTAACTCTCGCTGCTGTGGGCC	475			
Db					
QY	815 CGTCTCTCAAGTGGGTGAAGCAAAACATGCACGACTCATCTGGCTGGCCCTGTGGCTG	874			
Db					
QY	476 TTAGTGAACCTTTAGAACTTCTACCCATGCTCACTTTTCATCAATTTCTGTCCTCCAAAGAG	535			
Db					
QY	875 GTTGTCAACATCTGCCTCATCCCAACCTCATCTTCTGTCACCAACAGCTCCCAAGACAC	934			
Db					
QY	536 GGCAGTAACATGCAATGCACTATGCAAGTTCTGGAACCCCTGAACAAATCTATTACAGC	595			
Db					
QY	935 AGCACCTTTTGGCATGACACCAACCAACCTGAGGAGTTCCGACCACTACGTCGCTACAGC	994			
Db					
QY	596 CTCGTGCTGACITTTGTTGGCTTCTCTAAATCTCTCTCTCTCTCTCTCTCTCTCTCT	655			
Db					
QY	995 TCTCTCATGCTGCGCTCTCTCTGCTATTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCT	1054			
Db					
QY	656 AAGATGGTAGTCTTTTAAAGAGGAGGAGC-----CAGCAGCAAGCAACTGCCCTGC	707			
Db					
QY	1055 CTGATGGCCAAGAGGCTCTGCAAGCGCAGCTCTCCCAAGCCCGCCCGCTGCCCTCC	1114			
Db					
QY	708 CACTGGACAAACCCCAAC--GCCTGGTGGHCCCTGGGGTGTGATCTTCTCTATATCTTTC	766			
Db					
QY	1115 TACAAGAGCGTCCATCAAGATGATCATCTGCTACCGCTCTCGCCATTTTGCCTTT	1174			
Db					
QY	767 ACACCTTATCATATCATGCGCAATTTGAGGATCGCTCACGCTGGATAGTTGGCCACAA	826			
Db					
QY	1175 GTGCCCTTCCACATCACCGCGACCTCTACTACACTCTCGC---TACTTCCAAGCTGAT	1231			
Db					
QY	827 GGAATGACAGAGGCGCATCAATCTATATACACATGACACGCGCTCTGGCCCTTCTG	886			
Db					
QY	1232 TGTACAGCCTCAACATCATCAACTTCACTTACAAGATCACGCGACCCCTGGCAGCATC	1291			
Db					
QY	887 AACAGTGCCATCAATCCCATCTTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	946			
Db					
QY	1292 AACAGTGCTGGACCCCATCTGTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1351			
Db					
RESULT 14	AC101335	69462 bp	DNA	linear	HTG 23-NOV-2001
LOCUS	Mus musculus clone RP23-109K24, LOW-PASS SEQUENCE SAMPLING.				
DEFINITION	AC101335				
ACCESSION	AC101335.1	GI:17060110			
VERSION	HTG: HTGS_PHASE0.				
KEYWORDS	Mus musculus.				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 69462)				
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.				
TITLE	Unpublished				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 69462)				
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhagalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McSheeters, R., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,				

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L16470  
Center clone name: 109\_K\_24

\* NOTE: This record contains 85 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
\* 705 804 : contig of 704 bp in length  
\* 805 1533 : contig of 729 bp in length  
\* 1534 1633 : gap of 100 bp  
\* 1634 2366 : contig of 733 bp in length  
\* 2367 2466 : gap of 100 bp  
\* 2467 3194 : contig of 728 bp in length  
\* 3195 3294 : gap of 100 bp  
\* 3295 3987 : contig of 693 bp in length  
\* 3988 4087 : gap of 100 bp  
\* 4088 4826 : contig of 739 bp in length  
\* 4827 4926 : gap of 100 bp  
\* 4927 5647 : contig of 721 bp in length  
\* 5648 5747 : gap of 100 bp  
\* 5748 6477 : contig of 730 bp in length  
\* 6478 6577 : gap of 100 bp  
\* 6578 7292 : contig of 715 bp in length  
\* 7293 7392 : gap of 100 bp  
\* 7393 8113 : contig of 721 bp in length  
\* 8114 8213 : gap of 100 bp  
\* 8214 8933 : contig of 720 bp in length  
\* 8934 9033 : gap of 100 bp  
\* 9034 9758 : contig of 725 bp in length  
\* 9759 9858 : gap of 100 bp  
\* 9859 10651 : contig of 693 bp in length  
\* 10652 10651 : gap of 100 bp  
\* 10652 11386 : contig of 735 bp in length  
\* 11387 11486 : gap of 100 bp  
\* 11487 12223 : contig of 737 bp in length  
\* 12224 12323 : gap of 100 bp  
\* 12324 13057 : contig of 734 bp in length  
\* 13058 13157 : gap of 100 bp  
\* 13158 13845 : contig of 688 bp in length  
\* 13846 13945 : gap of 100 bp  
\* 13946 14669 : contig of 724 bp in length  
\* 14670 14769 : gap of 100 bp  
\* 14770 15482 : contig of 713 bp in length  
\* 15483 15582 : gap of 100 bp  
\* 15583 16293 : contig of 711 bp in length  
\* 16294 16393 : gap of 100 bp  
\* 16394 17089 : contig of 696 bp in length  
\* 17090 17189 : gap of 100 bp  
\* 17190 17914 : contig of 725 bp in length  
\* 17915 18014 : gap of 100 bp  
\* 18015 18576 : contig of 562 bp in length  
\* 18577 18676 : gap of 100 bp  
\* 18677 19389 : contig of 713 bp in length  
\* 19390 19489 : gap of 100 bp  
\* 19490 20233 : contig of 744 bp in length  
\* 20234 20333 : gap of 100 bp  
\* 20334 21062 : contig of 729 bp in length  
\* 21063 21162 : gap of 100 bp  
\* 21163 21865 : contig of 703 bp in length  
\* 21866 21965 : gap of 100 bp  
\* 21966 22694 : contig of 729 bp in length  
\* 22695 22794 : gap of 100 bp  
\* 22795 23520 : contig of 726 bp in length  
\* 23521 23620 : gap of 100 bp  
\* 23621 24337 : contig of 717 bp in length  
\* 24338 24437 : gap of 100 bp  
\* 24438 25149 : contig of 712 bp in length  
\* 25150 25249 : gap of 100 bp  
\* 25250 25905 : contig of 656 bp in length  
\* 25906 26005 : gap of 100 bp  
\* 26006 26738 : contig of 733 bp in length  
\* 26739 26838 : gap of 100 bp  
\* 26839 27572 : contig of 734 bp in length  
\* 27573 27672 : gap of 100 bp  
\* 27673 28400 : contig of 728 bp in length  
\* 28401 28500 : gap of 100 bp  
\* 28501 29196 : contig of 696 bp in length  
\* 29197 29296 : gap of 100 bp  
\* 29297 30013 : contig of 717 bp in length  
\* 30014 30113 : gap of 100 bp  
\* 30114 30821 : contig of 708 bp in length  
\* 30822 30921 : gap of 100 bp  
\* 30922 31642 : contig of 721 bp in length  
\* 31643 31742 : gap of 100 bp  
\* 31743 32459 : contig of 717 bp in length  
\* 32460 32559 : gap of 100 bp  
\* 32560 33291 : contig of 732 bp in length  
\* 33292 33391 : gap of 100 bp  
\* 33392 34135 : contig of 744 bp in length  
\* 34136 34235 : gap of 100 bp  
\* 34236 34946 : contig of 711 bp in length  
\* 34947 35046 : gap of 100 bp  
\* 35047 35775 : contig of 729 bp in length  
\* 35776 35875 : gap of 100 bp  
\* 35876 36605 : contig of 730 bp in length  
\* 36606 36705 : gap of 100 bp  
\* 36706 37439 : contig of 734 bp in length  
\* 37440 37539 : gap of 100 bp  
\* 37540 38270 : contig of 731 bp in length  
\* 38271 38370 : gap of 100 bp  
\* 38371 39096 : contig of 726 bp in length  
\* 39097 39196 : gap of 100 bp  
\* 39197 39917 : contig of 721 bp in length  
\* 39918 40017 : gap of 100 bp  
\* 40018 40726 : contig of 709 bp in length  
\* 40727 40826 : gap of 100 bp  
\* 40827 41544 : contig of 718 bp in length  
\* 41545 41644 : gap of 100 bp  
\* 41645 42332 : contig of 688 bp in length  
\* 42333 42432 : gap of 100 bp  
\* 42433 43168 : contig of 736 bp in length  
\* 43169 43268 : gap of 100 bp  
\* 43269 44006 : contig of 738 bp in length  
\* 44007 44106 : gap of 100 bp  
\* 44107 44838 : contig of 732 bp in length  
\* 44839 44938 : gap of 100 bp  
\* 44939 45673 : contig of 735 bp in length  
\* 45674 45773 : gap of 100 bp  
\* 45774 46500 : contig of 727 bp in length  
\* 46501 46600 : gap of 100 bp  
\* 46601 47316 : contig of 716 bp in length  
\* 47317 47416 : gap of 100 bp  
\* 47417 48119 : contig of 703 bp in length  
\* 48120 48219 : gap of 100 bp







GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2003, 10:58:41 ; Search time 393 seconds  
(without alignments)  
8841.815 Million cell updates/sec

Title: us-09-891-138a-1

Perfect-score: 1543  
Sequence: 1 gcccctggcagagtttctg.....tgcctaaataaatcaatata 1543

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :				
1:	1	1	1	1
2:	2	2	2	2
3:	3	3	3	3
4:	4	4	4	4
5:	5	5	5	5
6:	6	6	6	6
7:	7	7	7	7
8:	8	8	8	8
9:	9	9	9	9
10:	10	10	10	10
11:	11	11	11	11
12:	12	12	12	12
13:	13	13	13	13
14:	14	14	14	14
15:	15	15	15	15
16:	16	16	16	16
17:	17	17	17	17
18:	18	18	18	18
19:	19	19	19	19
20:	20	20	20	20
21:	21	21	21	21
22:	22	22	22	22
23:	23	23	23	23
24:	24	24	24	24

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1543	100.0	1543	24	ABK12957
2	592.4	38.4	1005	21	AA001135
3	592.4	38.4	1005	21	AA046036
4	592.4	38.4	1436	24	AB190790
5	592.4	38.4	1542	24	AA024958
6	589.2	38.2	1996	18	AA071900
7	587.2	38.1	1428	18	AA075146
8	580.4	37.6	1385	21	AA081122
9	126.6	8.2	1011	24	AA043942

10	126.6	8.2	1014	22	AA07948	Human cDNA encoding
11	126.6	8.2	1014	24	ABN5630	Human P2Y-like rec
12	126.6	8.2	1014	24	ABK11381	Human DNA encoding
13	126.6	8.2	1014	24	AAK98323	Human purinergic-r
14	126.6	8.2	1288	24	ABL56197	Human P2Y1-li enco
15	126.6	8.2	1288	24	AA08362	Human cDNA encoding
16	126.6	8.2	1729	23	ABV24026	Human prostate exp
17	126.6	8.2	1729	23	ABV25767	Human prostate exp
18	126.6	8.2	1729	23	ABV29909	Human prostate exp
19	126.6	8.2	1729	23	ABV30024	Human prostate exp
20	126.6	8.2	5435	24	ABL56198	Human P2Y1-li enco
21	126.6	8.2	9905	24	AAK98324	Human purinergic-r
22	125	8.1	1014	24	ABQ78847	Human G-protein co
23	125	8.1	1014	24	AA034278	Human G-protein co
24	125	8.1	1014	24	AA026370	Human AXOR89 (g-pr
25	119.2	7.7	1020	22	AAH51011	Human G-protein co
26	119	7.7	1313	22	AAK52430	Human ngPCR54 codi
27	112.6	7.3	740	23	ABV15662	Human polynucleoti
28	104	6.7	1076	24	AA029667	Human G-protein co
29	103.2	6.7	6721	24	AA018600	Purinergic recepto
30	100.8	6.5	6721	24	AA018599	Human G-protein co
31	97.8	6.3	1002	22	AAH73516	Human purinergic r
32	97.8	6.3	3143	24	AA017746	Human P2Y-like rec
33	96	6.2	831	24	ABN85629	Human genomic clon
34	90	5.8	539	23	ABV39127	Human P2Y-like rec
35	90	5.8	539	23	ABV45465	Human prostate exp
36	90	5.8	1063	24	AA098135	Human prostate exp
37	88.4	5.7	1429	18	AA074321	Human DNA for pote
38	87	5.6	1163	22	AA04981	Human P2Y4 recepto
39	85.4	5.5	2025	22	AA04980	Human purinergic r
40	84.8	5.5	2796	24	ABQ91998	Human purinergic r
41	83.8	5.4	1113	23	AB198010	Human NF-kB activa
42	83	5.4	850	24	AA017747	Non-endogenous hum
43	83	5.4	2138	24	AA032937	Human P2Y-like G p
44	83	5.4	2138	24	AA032934	Mus musculus GPCR
45	80	5.2	1068	23	AB197979	Mus musculus GPCR

## ALIGNMENTS

RESULT 1  
ABK12957  
ID ABK12957 standard; DNA; 1543 bp.  
XX  
AC ABK12957;  
XX

DT 09-APR-2002 (first entry)

DE DNA sequence of mouse G-protein coupled receptor TGR18 gene.

Mouse; G-protein coupled; receptor; GPCR; TGR18; kidney disease;  
signal transduction modulator; cerebral cavernous malformation;  
hyperlipidemia; obesity; dyslexia; cardiac myxoma; renal failure;  
nephritis; hypertension; liver disease; cirrhosis; blood disorder;  
spleen-associated disorder; immune disorder; gene; ds.

OS Mus sp.

FH Key Location/Qualifiers

FT CDS 44..997

FT /\*tag= a

FT /product= "Mouse G-protein coupled receptor TGR18"

XX WO200200719-A2.

XX 03-JAN-2002.

XX 25-JUN-2001; 2001WO-US20363.

XX 23-JUN-2000; 2000US-213461P.

XX (TULA-) TULARIK INC.

Lin DC, Zhao J, Chen J, Cutler G;  
WPI: 2002-147880/19.  
P-PSDB; AAU74904.

New G-protein coupled receptor polypeptides, useful for identifying modulators of signal transduction for treating kidney disease, hyperlipidemia, obesity, dyslexia and cardiac myxoma -

Claim 18; Page 58; 78pp; English.

The present invention relates to a new G-protein coupled receptor (GPCR) polypeptide comprising greater than 70% amino acid sequence identity to the amino acid sequence of human GPCRs TGR62, TGR21, TGR130.1, TGR130.2, human TGR213 or TGR92, 80% amino acid sequence identity to mouse TGR18 or 90% amino acid sequence identity to human novel edg receptor protein, as defined in the specification. The GPCR covalently linked to a solid phase is useful for identifying a compound that modulates signal transduction. The identified compounds are useful for treating kidney disease, cerebral cavernous malformations, hyperlipidemia, obesity, dyslexia and cardiac myxoma. The molecules of the invention are useful for diagnosing disorders or conditions such as kidney-related conditions or diseases such as renal failure, nephritis, nephrotic syndrome, asymptomatic urinary abnormalities, renal tubule defects, hypertension and nephrolithiasis, liver-related disease or condition e.g. cirrhosis, infiltrations, lesions, functional disorders and jaundice and spleen-associated disorders or conditions e.g. splenic enlargement, immune disorders, blood disorders and others. Modulation of the polypeptide of the invention is useful to treat or prevent any of the above conditions or diseases. The present nucleic acid sequence encodes the mouse GPCR TGR18 protein of the invention. This sequence encodes one of seven novel G protein coupled receptors of the invention (ABK12957-ABK12964).

Sequence 1543 BP; 438 A; 352 C; 293 G; 460 T; 0 other;  
Query Match 100.0%; Score 1543; DB 24; Length 1543;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	GCTCTGCGCAGAGTTTCTCTGCGACAGAAAGCGGACGACGAGAAATGGCAGCAAAATTATC	60
1	GCTCTGCGCAGAGTTTCTGTCGACAGAAAGCGGACGACGAGAAATGGCAGCAAAATTATC	60
61	TTGTGAGAAATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACTCTCTGCATTTTA	120
61	TTGTGAGAAATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACTCTCTGCATTTTA	120
121	TGCAATCGAGTTCATTTTGGACATGCTTTGGGAATGTCACHTGTGGTTCGGCTACCTCTT	180
121	TGCAATCGAGTTCATTTTGGACATGCTTTGGGAATGTCACHTGTGGTTCGGCTACCTCTT	180
181	CTGCATGAAGAACTGGACACGCAATGCTATCTTTTAACTTTCCATCTCTGACCTT	240
181	CTGCATGAAGAACTGGACACGCAATGCTATCTTTTAACTTTCCATCTCTGACCTT	240
241	TGCTTTCTCTGTGCACCCCTTCCCATCTCTGATAAAGAGTTATGCCAATGATGAAGGGACCTA	300
241	TGCTTTCTCTGTGCACCCCTTCCCATCTCTGATAAAGAGTTATGCCAATGATGAAGGGACCTA	300
301	TGGAGATGTTCTCTGTATAACCAACCGATATGTCCTTACACCAACCTCTACACCAAGCAT	360
301	TGGAGATGTTCTCTGTATAAAGCAACCGATATGTCCTTACACCAACCTCTACACCAAGCAT	360
361	CTCTTCTCTCACTTTTCATTAGCATGGACCGGATATCTGCTCATGAAGTACCTTTCCGAGA	420
361	CTCTTCTCTCACTTTTCATTAGCATGGACCGGATATCTGCTCATGAAGTACCTTTCCGAGA	420
421	ACACTTTCTTACAAAAGAGGAATTTGCCATTTTAACTCGCTGCTCTGGGCTTAGT	480
421	ACACTTTCTTACAAAAGAGGAATTTGCCATTTTAACTCGCTGCTCTGGGCTTAGT	480

481	QY	GACCTTAGAAGTTCTTACCATGCTCACTTTTCATCAATTCTGTCCCAAAGAAGAGGGCAG	540
481	DB	GACCTTAGAAGTTCTTACCATGCTCACTTTTCATCAATTCTGTCCCAAAGAAGAGGGCAG	540
541	QY	TAACTGCATCGACTATCGCAGTTCTGGAACCCCTGAACAACATCTCATTTACAGCCTCTG	600
541	DB	TAACTGCATCGACTATCGCAGTTCTGGAACCCCTGAACAACATCTCATTTACAGCCTCTG	600
601	QY	CCTGACCTTTGTTGGGCTTCCTAAATCCTCTCTCTGTGTGATGTCTTCTTCTACTACAAGAT	660
601	DB	CCTGACCTTTGTTGGGCTTCCTAAATCCTCTCTCTGTGTGATGTCTTCTTCTACTACAAGAT	660
661	QY	GGTAGTCTCTTTAAAGAGGAGGACGACGACGAACCACTGCCCTGCCACTGGACAAAACC	720
661	DB	GGTAGTCTCTTTAAAGAGGAGGACGACGACGAACCACTGCCCTGCCACTGGACAAAACC	720
721	QY	CCAGCGCTGGTGCTCTGCCGTTGTGATCTTCTATACTCTTTCACACCCCTATCATAT	780
721	DB	CCAGCGCTGGTGCTCTGCCGTTGTGATCTTCTATACTCTTTCACACCCCTATCATAT	780
781	QY	CATGCGCAAAATTGAGGATCGCCTCACGCCCTGGGATAGTTGGCCNCAAGGATGTACAGAA	840
781	DB	CATGCGCAAAATTGAGGATCGCCTCACGCCCTGGGATAGTTGGCCNCAAGGATGTACAGAA	840
841	QY	GGCCATCAAACTATATACACACTGACACGGCCTCTGGCCCTTCTGAAACAGTGCCTCAA	900
841	DB	GGCCATCAAACTATATACACACTGACACGGCCTCTGGCCCTTCTGAAACAGTGCCTCAA	900
901	QY	TCCCATCTTCTACTTCCTTCATGGGAGACCAATACAGAGATGCTGATTTAGTAAGTTCA	960
901	DB	TCCCATCTTCTACTTCCTTCATGGGAGACCAATACAGAGATGCTGATTTAGTAAGTTCA	960
961	QY	ACAATACTTCAAGTCCCTTACATCCTTCAGACATGAGCTGCTGGAATGCAAGTCTTCACT	1020
961	DB	ACAATACTTCAAGTCCCTTACATCCTTCAGACATGAGCTGCTGGAATGCAAGTCTTCACT	1020
1021	QY	CAGCCAAAATGAGACACTTGATAAACAGTCTCTGTCAGTTGAGTTTTAAGTAAGTAACC	1080
1021	DB	CAGCCAAAATGAGACACTTGATAAACAGTCTCTGTCAGTTGAGTTTTAAGTAAGTAACC	1080
1081	QY	ACATTTCTAGGCTTTAGCTTTCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG	1140
1081	DB	ACATTTCTAGGCTTTAGCTTTCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG	1140
1141	QY	GTTCACATGAATCAGAAGCAGCTCTCTGTTCTTGATTTTAGGTTAPACCCAGAGTATGG	1200
1141	DB	GTTCACATGAATCAGAAGCAGCTCTCTGTTCTTGATTTTAGGTTAPACCCAGAGTATGG	1200
1201	QY	AAAAATAAGGCATGAGAAAGCATTGACATCTTCACITTAGAACTGNAACAAAAGAGAACA	1260
1201	DB	AAAAATAAGGCATGAGAAAGCATTGACATCTTCACITTAGAACTGNAACAAAAGAGAACA	1260
1261	QY	AATATTTGCAATGTTTGGACACTTAGGATCTGAAATCTTGGAATTTTAAGACCTCTTTT	1320
1261	DB	AATATTTGCAATGTTTGGACACTTAGGATCTGAAATCTTGGAATTTTAAGACCTCTTTT	1320
1321	QY	TCATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAAATGCPGAATGCATTTTCATCAT	1380
1321	DB	TCATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAAATGCTGAATGCATTTTCATCAT	1380
1381	QY	TGTCAGGTCGATAGACGTTTCTGGAATAGTCTTATTTTATTTCTGTTAAATATAAAA	1440
1381	DB	TGTCAGGTCGATAGACGTTTCTGGAATAGTCTTATTTTATTTCTGTTAAATATAAAA	1440
1441	QY	TTTATCTGGAANAATGAATAATAATTCAAATGTAACAATAGATTTTCTTATTTGGAATTTAT	1500
1441	DB	TTTATCTGGAANAATGAATAATAATTCAAATGTAACAATAGATTTTCTTATTTGGAATTTAT	1500
1501	QY	ATTTCTTGAAAAATAACTGCTGCTCCCTAAATAAATAAATAAATA	1543
1501	DB	ATTTCTTGAAAAATAACTGCTGCTCCCTAAATAAATAAATAAATA	1543

## RESULT 2

AD01135  
 ID AAD01135 standard; cDNA: 1005 BP.  
 AC AAD01135;  
 XX  
 XX  
 XX 02-NOV-2000 (first entry)  
 XX  
 XX Human orphan G protein-coupled receptor hCHN10 cDNA.  
 DE  
 XX  
 XX Human; orphan G protein-coupled receptor; GPCR; hCHN10; drug screening;  
 KW transmembrane receptor; expressed sequence tag; EST; signal cascade; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 1..1005  
 FT CDS /\*tag= a  
 FT /product= "hCHN10"  
 FT /note= "Human orphan G protein-coupled receptor"  
 XX  
 XX WO200031258-A2.  
 PN  
 XX  
 XX 02-JUN-2000.  
 PD  
 XX  
 XX 13-OCT-1999; 99WO-US23687.  
 PF  
 XX  
 XX 20-NOV-1998; 98US-0109213.  
 PR 16-FEB-1999; 99US-0120416.  
 PR 26-FEB-1999; 99US-0121852.  
 PR 12-MAR-1999; 99US-0123946.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0136567.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 29-JUN-1999; 99US-0144448.  
 PR 29-SEP-1999; 99US-0156555.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156634.  
 PR 01-OCT-1999; 99US-0157280.  
 PR 01-OCT-1999; 99US-0157281.  
 PR 01-OCT-1999; 99US-0157282.  
 PR 01-OCT-1999; 99US-0157293.  
 PR 01-OCT-1999; 99US-0157294.  
 PR 12-OCT-1999; 99US-0416760.  
 PR 12-OCT-1999; 99US-0417044.  
 XX  
 XX (AREN-) ARENA PHARM INC.  
 PA  
 XX  
 XX Chen R, Dang HT, Liaw CW, Lin I;  
 PI WPI; 2000-400068/34.  
 DR P-PSDB; AAY71308.  
 DR  
 XX  
 XX Novel human orphan G protein-coupled receptors and the encoding cDNAs  
 PT for use in the identification of G protein-coupled receptor agonists -  
 XX  
 XX Claim 69; Page 86; 102pp; English.  
 PS  
 XX  
 XX The present sequence is a cDNA encoding hCHN10, an endogenous human  
 CC orphan G protein-coupled receptor (GPCR), expressed in kidney and  
 CC thyroid. The hCHN10 cDNA was identified using the human EST (expressed  
 CC sequence tag) 1365839 as a probe. The orphan GPCR of the invention, like  
 CC all GPCRs has seven transmembrane alpha helices with an extracellular  
 CC N-terminus and an intracellular C-terminus. However, no endogenous  
 CC ligands has yet been identified for the proteins of the invention. The  
 CC orphan GPCRs may be used in the identification of their endogenous  
 CC ligands, and to screen potential GPCR agonists and antagonists for use as  
 CC pharmaceutical agents. The proteins may also be used in the study of

CC GPCR-mediated signalling cascades, and to elucidate their precise role in  
 CC normal and diseased human conditions. Nucleic acid encoding human orphan  
 CC GPCRs may be used for tissue localisation expression analysis to provide  
 CC information about their function in healthy and pathological states.  
 XX  
 XX  
 SQ Sequence 1005 BP; 248 A; 236 C; 196 G; 325 T; 0 other;

Query Match 38.4%; Score 592.4; DB 21; Length 1005;  
 Best Local Similarity 75.5%; Pred. No. 1e-139;  
 Matches 750; Conservative 0; Mismatches 241; Indels 3; Gaps 1;  
 QY 39 GCAGAATGGCACAGAATTTATCTTGAGAAATTTGGTTGGCAACAGAGGCTATCTTGAATA 98  
 DB 8 GGATCAATGCGAATGCAACITGCAAAAACCTGGCTGCAGCAGGCTGCCCTGGAAA 67  
 QY 99 AGTACTACCTCTCTGCAATTTTATGCAATCGAGTTTCATTTTGGACTGCTTGGGAATGTCA 158  
 DB 68 AGTACTACCTTTCCATTTTATGGGATTTGAGTTCTGTTGGGAGTCTCTTGAATAACCA 127  
 QY 159 CTGTGGTGTTCGGCTACCTCTTCTGCAAGAACTGGAACAGCAATGCTATCTTT 218  
 DB 128 TTGTTGTTTACGGCTACATCTTCTCTGGAAGAACTGGAACAGCAATATTTATCTCT 187  
 QY 219 TTAACCTTTCCATCTCTGACTTTTGCTTCTGTCACCTTCCCATCTCTGATAAAGATT 278  
 DB 188 TTAACCTCTCTCTGACTTGTCTGTCACCTTCTCTGTCACCTTCCCATCTCTGATAAAGATT 247  
 QY 279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGATAAGCAACGATATGCTTC 338  
 DB 248 ATGCCAATGGAACCTGGATATATGGAGAGCTGCTCTGCAATGAACCAACGATATGCTTC 307  
 QY 339 ACACCAACCTCTACACGACGATCCTCTTCCCTTCTTCTTATGAGTATGAGCAGGATATCGC 398  
 DB 308 ATGCCAACCTCTATACCAGCATCTCTTCTTCTTCTTATGAGTATGAGCAGGATATCGC 367  
 QY 399 TCATGAAGTACCTTTCCGAGAACACTTTCTACAAAAGAGAAATTTGCCATTTTAACTCT 458  
 DB 368 TAAATTAAGTATCCTTTCCGAGAACACTTTCTGCAAAAAGAGAAATTTGCCATTTTAACTCT 427  
 QY 459 CGCTGCTGCTCTGGGCTTGTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAAT 518  
 DB 428 CTTGGCCATTTGGGCTTTTGTAGTACCTTAGAGTTTACTACCCATCTTCCCTTTAAATC 487  
 QY 519 CTGTCCCAAGAGAGGCGCAGTAACTGCATCGACTATGCAAGTTCTGGAACCTCAAC 578  
 DB 488 CTGTTTAACTGACAAATGCGACCACTGTATGATTTTGCAGTTCTGGAGACCCCACT 547  
 QY 579 ACAATCTCATTTACAGCCTCTGCTGACTTTGTTGGGCTTCTTAACTCTCTCTCTGTA 638  
 DB 548 ACAACCTCATTTACAGCATGCTCTACACACTGTTGGGCTTCTTATCTCTCTTTTGTGA 607  
 QY 639 TGTGCTTCTCTACTACAGAGTGTAGTCTTCTTAAAGAGAGGAGCCAGCAGCAACAA 698  
 DB 608 TGTGTTTCTTTTATACAGATGCTCTCTTCTTCTTAAAGAGAGGAGATAGCAGGTTGCTA 667  
 QY 699 CTGCCCTGCCACTGGACAACCCCAAGCCCTGGTGGCTCTGGGCTTGTGATCTTCTCTA 758  
 DB 668 CTGCTCTGCCCTTGGAAAGCCTCTCAACTTGGTCACTATGCGAGTGGTATCTCTCTG 727  
 QY 759 TACTCTTCACACCTTATCATATCATGCGCAATTTGAGGATCGGCTCACGCTCGATAGTT 818  
 DB 728 TGTCTTTTACACCTTATCAGTCTCATGCGGAATGTGAGGATCGCTTACGCTGGGAGTT 787  
 QY 819 G---GCCAAGAGTGTACAGAAAGCCCATCAATCATATATACACTACACAGGCTC 875  
 DB 788 GGAAGCAGTATCAGTGCACCTAGGTCGTCATCAACTCTCTTTTATCTGTACACAGGCTT 847  
 QY 876 TGGCCTTTCTGAACAGTGGCCATCAATCCCATCTTCTTCTTCTCTCATGGAGACCATACA 935  
 DB 848 TGGCTTTCTGAACAGTGTCAATCAACCTCTCTTCTTCTTCTTCTTGGGAGATCACTTCA 907  
 QY 936 GAGAGATGCTTATAGTAAAGTTTACAGAAATPACITTCAGTCCCTTACATCTCTCAGGACAT 995

Db 908 GGCACATGCTGATGATCAATCAACTGACAGACACAACTTCAATCCCTTACATCCTTTAGCAGAT 967  
 QY 996 GAGCTGCTGGATGAGGCTTCTTCACTCAGCCAAA 1029  
 Db 968 GGGCTCATGAACCTCCTACTTCTTCAATTCAGAGAAA 1001

RESULT 3  
 AAA46036  
 ID AAA46036 standard; cDNA; 1005 BP.

AC AAA46036;  
 XX 22-AUG-2000 (first entry)  
 XX Human G protein coupled receptor hCHN10 encoding cDNA SEQ ID NO:37.  
 DE Human; G protein coupled receptor; GPCR; transmembrane receptor;  
 KW identification; agonist; screening; therapeutic; pharmaceutical;  
 KW mutant; ss.  
 XX Homo sapiens.

OS WO200022131-A2.  
 XX 20-APR-2000.  
 XX 13-OCT-1999; 99WO-US24065.  
 XX 13-OCT-1998; 98US-0170496.  
 PR 12-NOV-1998; 98US-018029.  
 PR 20-NOV-1998; 98US-0109213.  
 PR 27-NOV-1998; 98US-0110060.  
 PR 16-FEB-1999; 99US-0120416.  
 PR 26-FEB-1999; 99US-0121852.  
 PR 12-MAR-1999; 99US-0123944.  
 PR 12-MAR-1999; 99US-0123945.  
 PR 12-MAR-1999; 99US-0123946.  
 PR 12-MAR-1999; 99US-0123948.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 12-MAR-1999; 99US-0123951.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 28-MAY-1999; 99US-0137567.  
 PR 30-JUN-1999; 99US-0141448.  
 PR 27-AUG-1999; 99US-0151114.  
 PR 03-SEP-1999; 99US-0152524.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156555.  
 PR 29-SEP-1999; 99US-0156634.  
 XX (AREN-) ARENA PHARM INC.

XX Behan DP, Lehmann-Bruinsma K, Chalmers DF, Chen R, Dang HT;  
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;  
 XX WPI; 2000-317986/27.  
 DR P-PSDB; AAB02842.  
 XX Non-endogenous, human G protein-coupled receptors for screening  
 PT receptor, inverse or partial agonists useful as therapeutic agents  
 PT Example 1; Page 116; 187pp; English.  
 PS The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCR), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically, the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,

CC inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX Sequence 1005 BP; 248 A; 236 C; 196 G; 325 T; 0 other;  
 SQ Query Match 38.4%; Score 592.4; DB 21; Length 1005;  
 Best Local Similarity 75.5%; Pred. No. le-139;  
 Matches 750; Conservative 0; Mismatches 241; Indels 3; Gaps 1;  
 QY 39 GCAGATGGCAGCAATTTATCTTGTGAGAAATGTTGGCAACAGAGGCTATCTTGAATA 98  
 Db 8 GGATCATGGCATGGAATGCAACTTGCAGAAAGAGGCTGGCAGAGAGGCTGCCCTGGAAA 67  
 QY 99 AGTACTACTCTCTGCAATTTATGCAATCGAGTTCAATTTTGGAGCTGCTGGGAATGTCA 158  
 Db 68 AGTACTACTCTTCCATTTTATGGGATGAGTTGCTGTGGAGTCTCTTGAATATACCA 127  
 QY 159 CTGCGGTGTTGGCTACCTCTCTGCATGAAGAACTGGAACAGCAGCAATCTCTATCTTT 218  
 Db 128 TTGTTGTTTACGGCTACATCTTCTCTGAAGAACTGGAACAGCAGTAATATTATCTCT 187  
 QY 219 TTAACCTTTCCATCTCTGACTTTCTCTGTGACCCCTTCCCATCTCTGATAAGAGTT 278  
 Db 188 TTAACCTCTCTGCTCTGACTTAGCTTTCTGTGACCCCTCCCATGCTGATAAGAGTT 247  
 QY 279 ATGCCAATGATAAGGGACCTATGAGATGTTCTCTGTATAGCAACCGATATGTCTTC 338  
 Db 248 ATGCCAATGGAACCTGGATATATGAGAGCTGCTCTGATAGCAACCGATATGTCTTC 307  
 QY 339 ACACCAACCTCTACACAGCAGCTCTCTCCACCTTTCATTAGCATGACCGATATCTGC 398  
 Db 308 ATGCCAACCTCTATACAGCAGCTCTCTCTACCTTTTATCAGCATAGATGATCTTGA 367  
 QY 399 TCATGAAGTACCTTTTCCGAGAACACTTTCTACAAAGAGAAATTTGCCATTTAATCT 458  
 Db 368 TAATTAAGTATCTTCCGAGAACCTTCTGCAAGAAAGAGTTTCTATTTAATCT 427  
 QY 459 CGCTGGCTGCTGGGCTTCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 518  
 Db 428 CTTGGCCATTTGGGTTTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 487  
 QY 519 CTGTCCCAAGAGAGGAGGAGTAACTGATCGACTATGCAAGTCTTGGAAACCTGTGAC 578  
 Db 488 CTGTATTAATGACATGACATGACACCTGTAATGATTTTGAAGTTCTGGAGAGCTT 547  
 QY 579 ACAATCTCATTTACAGCTCTGCTGACTTTCTGGGCTTCTTAATTTCTCTCTCTGTA 638  
 Db 548 ACAACCTCATTTACAGCATGTGTCTAACACTGTGGGGTCTCTTATCTCTTTTGTGA 607  
 QY 639 TGTGCTTCTTCTACTACAAGATGGTGTCTTCTTAAAGAGAGGAGGCCAGCAAGCAA 698  
 Db 608 TGIGTTTCTTTTATTAAGATGCTCTCTCTCTAAAGAGAGGAATAGGAGGTTGCTA 667  
 QY 699 CTGCCCTGGCAGTGGACAAACCCCAAGCTGCTGGTGTCTGGGTTGTGTGATCTCTCTA 758  
 Db 668 CTGCTGCGCCCTTGAAGAGCTCTCAACTTGTGCTATCATGGAGTGGTAACTCTCTCTG 727  
 QY 759 TACTCTTCACACCTATCATATCATGCAATTTGAGGATCGCTCAGCCCTGGATAGTT 818  
 Db 728 TGCTTTTACACCTTATCAGCTCATCGGAAATGAGGATCGCTTACACGCTGGGAGTT 787  
 QY 819 G---GCCACAAGGATGTACACAGAAGGCCCAATTAATCTATATACACTGACACGCTTC 875  
 Db 788 GGAAGCAGTATCAGTGCACCTCAGGCTGCTCATCACTTCTTACATTTGTGACACGCTT 847  
 QY 876 TGGCCTTTCTGACAGCTGCCATCAATCCCATCTTCTTCTCTCTCATGGGAGCAATTACA 935  
 Db 848 TGGCCTTTCTGACAGCTGCCATCAATCCCATCTTCTTCTCTCTCATGGGAGCAATTACA 907  
 QY 936 GAGAGATGCTGATTAAGTGTACAGCAATTAATCTTCAAGTCCCTTACATCTCTTACAGCAT 995  
 Db 908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAATCCCTTACATCTCTTACAGAT 967







XX Human purinergic receptor P2U2 cDNA.  
 DE P2U2 receptor; purinergic receptor; diagnosis; therapy; ss.  
 KW Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 625..1629  
 FT /\*tag= a  
 FT  
 XX WC9720045-A2.  
 PN  
 XX 05-JUN-1997.  
 PD  
 XX 08-NOV-1996; 96WO-US18175.  
 XX  
 XX 15-NOV-1995; 95US-0559524.  
 PR 15-NOV-1995; 95US-0006782.  
 XX  
 XX (CORT-) COR THERAPEUTICS INC.  
 PA  
 XX Conley PB, Jantzen H;  
 PI  
 XX WPI: 1997-310601/28.  
 DR P-PSDB; AAW19854.  
 DR  
 XX  
 PT New isolated purinergic receptor sub-type - used to develop  
 PT products for diagnosis and therapy, e.g. for screening for agonists  
 PT and antagonists which can modulate activation  
 PT  
 XX Claim 3; Fig 1A-C; 36pp; English.  
 PS  
 XX A cDNA clone (AA171900) codes for a novel human purinergic receptor  
 CC subtype, designated P2U2 receptor (AAW19854), that is abundantly  
 CC expressed in kidney and in many cell lines of megakaryocytic or  
 CC erythroleukaemic origin and which is activated by ATP, UDP, UTP and  
 CC UDP. The clone was obtcd. by amplifying DAMI (ATCC CRL 9792) cell  
 CC cDNA using primers (see also AAT72104-05) based on transmembrane  
 CC regions of mouse P2u and chicken P2Y1 receptors, and use of the PCR  
 CC product to screen the DAMI cDNA library to isolate the full-length  
 CC clone. P2U2 nucleic acids can be used in the recombinant prodn. of  
 CC P2U2 receptor polypeptides and as probes.  
 XX  
 SQ Sequence 1996 BP; 513 A; 454 C; 381 G; 647 T; 1 other;

Query Match 38.2%; Score 589.2; DB 18; Length 1996;  
 Best Local Similarity 75.1%; Pred. No. 8.9e-139;  
 Matches 762; Conservative 0; Mismatches 248; Indels 4; Gaps 2;

QY 39 GCAGATGGCAGCAATTTATCTCTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATA 98  
 DB 632 GGATCATGGCATGGAATGCAACTTGCAGAAAGCTGCTGGCAGAGAGGCTGCCCTGGAAA 691  
 QY 99 AGTACTACTCTCTGCATTTATGCAATCGAGTTCATTTTGGAGTGTCTGGGAATGCA 158  
 DB 692 AGTACTACTCTCTGCATTTATGGAATGAGTGTGTTGGAGTGTCTGGGAATGCA 751  
 QY 159 CTGTGGTGTGGCTTACCTCTCTCTGATGAAAGAACTGGAACAGCAGCAATCTCTATCTTT 218  
 DB 752 TTGTTGTTTACGGTACATCTCTCTGAGAACTGGAACAGCAGCAATATTTATCTCTCT 811  
 QY 219 TTAACCTTTCACATCTCTGACATTTGCTTTCCTGTGACCCCTTCCCATCTGTATAAGAGTT 278  
 DB 812 TTAACCTCTCTCTCTGACATTTGCTTTCCTGTGACCCCTTCCCATCTGTATAAGAGTT 871  
 QY 279 ATGCCAATGATAAGGGGACCTATGGAGATGTCTCTGTATTAAGCAACCGATATGTGCTTC 338  
 DB 872 ATGCCAATGGAACCTGGATATATGAGACGTGCTCTGCAATGAACCAACCGATATGTGCTTC 931  
 QY 339 ACACCAACCTCTACACGAGCATCTCTCTCTCACTTTTCATAGCATGGACGATATCTGC 398  
 DB 932 ATGCCAACCTCTATACGAGCATCTCTCTCTCACTTTTATCAGCATAGATCGACTTGA 991

QY 399 TCATGAAGTACCCCTTTCCGAGAACACTTTCTACAAAAGAGGAATTTGCCATTTTAATCT 458  
 DB 992 TAATTAAGTATCCCTTTCCGAGAACACCCTTCTGCAAAAGAGAGAGTTTCTGCTATTTTAATCT 1051  
 QY 459 CGCTGGCTGTCTGGGCTTAGTGACCTTAGAGAGTTTCTACCCATGCTCACTTTTCATCAAT 518  
 DB 1052 CTTTGGCATTGGGTTTGTAGTAACCTTAGAGTTTCTACCCATGCTCCCTTATAAATC 1111  
 QY 519 CTGTCCCAAGAGAGGGCAGTAACATGCATCGACTATGCAAGTTTCTGAAACCCCTGAAC 578  
 DB 1112 CTGTTATAACTGACAAATGSCACCACCTCTAATGATTTTGGCAAGTTCTGGAGACCCCACT 1171  
 QY 579 ACAATCTCATTTACAGCCTCTGCTGACTTTTGGGCTTCTTAATCTCTCTCTGFGA 638  
 DB 1172 ACAACCTCATTTACAGCATGTCTTAACACTGTTGGGCTTCTTATCTCTCTTTTGTGA 1231  
 QY 639 TGTGCTTCTTCTACTACAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCA 698  
 DB 1232 TGTGTTCTTTTATACAGATGTCTCTTCTTAAAGCAGAGGAATAGGCGAGTTGCTA 1291  
 QY 699 CTGCCCTGCCACTGGACAAACCCCAAGCCTGGTGGTCTCTGGCGTTTGTGATCTTCTCTA 758  
 DB 1292 CTGCTCTGCCCTTGAAGGCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 1351  
 QY 759 TACTCTTCACACCCCTATCATATCATGCGCAATTTGAGGATCGCCTCAGCCTGATAGTT 818  
 DB 1352 TGTCTTTTACACCTTATCAGCTCATGGGAATGTGAGGATCGCTTACGCGCTGGGGAGTT 1411  
 QY 819 G---GCCAAGAGTGTACACAGAGGCGCATCAAAATCTATATACACACTGACACGCGCTC 875  
 DB 1412 GGAAGCAGTATCAGTGCACCTCAGGTGCTGTCATCACTCTTTTACATTTGACACGGGCTT 1471  
 QY 876 TGGCCTTTCTGAACAGTGCATCAATCCCATCTTCTTCTCTCATGCTGGGAGACCATTACA 935  
 DB 1472 TGGGCTTTCTGAACAGTGTCTCAACCCCTGTCTTCTTCTTCTTTTGGGAGATCACTTCA 1531  
 QY 936 GAGAGATGCTGATTAGTAAGTTTCAGACAATCACTTCAAGTCCCTTACATCTCTCAGGACAT 995  
 DB 1532 GGGACATGCTGATGAATCAACTGAGACACACACTTCAATCCCTTACATCTCTTAGCAGAT 1591  
 QY 996 GAGTGTGGATGAGGCTTCTTCACTCAGCCAAA--TGAGACACTTGTATAAACAG 1048  
 DB 1592 GGGCTCATGAACCTCTTCTTCTTCAATTCAGAGAAAAGTCAAGGGCTTGTGAAACAG 1645

RESULT 7  
 AAT75146  
 ID AAT75146 standard; cDNA; 1428 BP.  
 XX AC AAT75146;  
 XX XX  
 XX XX  
 XX 07-OCT-1997 (first entry)  
 XX Human ATP receptor cDNA.  
 DE ATP receptor; G-protein coupled receptor; agonist; antagonist; ss.  
 KW Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 92..1096  
 FT /\*tag= a  
 FT /transl\_except= (pos:725..727, aa:Ser)  
 FT /transl\_except= (pos:764..766, aa:Ser)  
 FT /transl\_except= (pos:820..822, Xaa)  
 FT /notes= "Xaa = unknown"  
 FT complement (92..100)  
 FT /\*tag= b  
 FT /note= "binding site for primer used to amplify  
 FT cDNA for baculovirus expression"  
 FT complement (92..109)  
 FT /\*tag= c



PF 06-APR-2000; 2000WO-0509070.  
 PR 09-APR-1999; 99US-0128695.  
 PR 14-JAN-2000; 2000US-0176052.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Ruben SM, Komatsoulis G;  
 PI WPI; 2000-619228/59.  
 XX P-PSDB; AAB45344.  
 DR New nucleic acid molecules encoding 49 human secreted proteins for  
 XX diagnosing, preventing, treating or ameliorating medical conditions and  
 XX used as food additives or preservatives -  
 XX Claim 1; Page 412; 454pp; English.  
 CC The polynucleotide sequences given in AAC81086 to AAC81134 encode the  
 CC human secreted proteins given in AAB45308 to AAB45356. AAB45357 to  
 CC AAB45384 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Examples of activities include:  
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;  
 CC cytotatic; cardiant; vasotropic; cerebrotective; nootropic;  
 CC neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 CC and vulnerary. The polynucleotides and polypeptides can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
 CC disorders, cerebrovascular disorders, angiogenesis, nervous system  
 CC disorders. The polypeptides caused by bacteria, viruses and fungi and ocular  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
 CC components. AAC81077 to AAC81085 and AAC81085 represent sequences used in  
 CC the exemplification of the present invention.  
 XX Sequence 1385 BP; 385 A; 296 C; 275 G; 429 T; 0 other;  
 SQ

Query Match 37.6%; Score 580.4; DB 21; Length 1385;  
 Best Local Similarity 75.2%; Pred. No. 1.3e-136;  
 Matches 763; Conservative 0; Mismatches 246; Indels 5; Gaps 3;

QY 39 GCAGAAATGGACAGAAATTTATCTGTGAGAAATTTGGTGGCAACAGAGGCTATCTTGAATA 98  
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 56 GGATCATGGCAGTGAATGCAACTTGCAGAAATGCTGGCAGCAGAGGCTCCCTGGNAA 115  
 QY 99 AGTACTACTCTCTGCATTTTATGCAATCGAGTTCAATTTTGGACTGCTTGGGAAATGCA 158  
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 116 AGTACTACTCTTCCATTTTATGGAATGAGTTGCTGTTGGGAGTCTCTGGAAATACCA 175  
 QY 159 CTGTGGTGTGGGTACCTCTCTGCAATGAGAACTGGAACAGCAGCAATCTCTATCTTT 218  
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 176 TTGTGTTTACGGCTACATCTCTCTGGAAGAACTGGAACAGCAGTAATATTTATCTCT 235  
 QY 219 TTAACCTTTTCCATCTCTGACTTTTCTTCTGTGACCCCTTCCCATCTCTGATAAAGAGTT 278  
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 236 TTAACCTCTCTGCTGACTTACTGTTTCTGTGACCCCTCCCATCTCTGATAAGAGTT 295  
 QY 279 ATGCCAATGATAAGGGACCTATGAGATGTTCTCTGTAAGCAACCGATATGTGCTTC 338  
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 296 ATGCCAATGGAACCTGATATGAGACGCTGCTCTCATGAAGCAACCGATATGTGCTTC 355  
 QY 339 ACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398  
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 356 ATGCCAACCTCTATACAGCAATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 415  
 QY 399 TCATGAAGTACCTTTTCCGAGAACACTTTCTACAAAAGAGAAATTTGCCAATTTTAAATCT 458  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 416 TAAATTAAGTATCTTTTCCGAGAACACCTTCTGCAAAAAGAGAGTTTCTATTTTAAATCT 475  
 QY 459 CGCTGGCTGTCTGGGCTTAGTGACCTTAGAGAGTTCTACCCATCTCTCTCTCTCTCTCTCTCT 518  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 476 CTTGGCCATTTGGGTTTATGACCTTAGAGTTTACTACCCATCTCTCTCTCTCTCTCTCTCTCT 535  
 QY 519 CTGTCCCCAAAAGAGAGGGCAGTAACTCATCGACTATGCAAGTTCTGGAACCCCTGAAC 578  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 536 CTGTTTATAACTGACAAATGGCACCCTCTGAATGATTTTGAAGTTCTGGAGACCCCAACT 595  
 QY 579 ACAATCTCAATTACAGCCTCTGCTGACTTTGTTGGGCTTCTCTAATCTCTCTCTCTCTCTCT 638  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 596 ACAACCTCATTTACAGCAATGTTCTTAACACTGTTGGGTTCTCTATTTCTCTCTCTCTCTCT 655  
 QY 639 TGTGCTTCT 698  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 656 TGTGTTTCTTTTATACAGATTTGCT 715  
 QY 699 CTGCCCTGCCACTGGACAAACCCCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 716 CTGCTCTGCCCTTGAAGGCTCTCAACTTGGTTCATCATGGCAGTGGTAACTCTCTCTCTCTCT 775  
 QY 759 TACTCTTACACCTTATCATATCATGCGCAATTTGAGGATGCTCTACGCTCTGATAGTT 818  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 776 TGTCTTTTACACCTTATCATGCTCATGGGATGTTGAGGATGCTCTACGCTCTGGGAGTT 835  
 QY 819 G---GCCACAAGGATGTACAGAAAGGCCCAATCAATCTATATACACTCACACGCGGCTC 875  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 836 GGAAGCAGTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894  
 QY 876 TGGCCTTTTGAACAGTGGCCATCAATCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 935  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 895 TGGCCTTTTGAACAGTGTCTATCAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 954  
 QY 936 GAGAGATGCTGATTAAGTTTACAGCAATCACTCAAGTCCCTTACATCTCTCTCTCTCTCTCTCT 995  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 955 GGGACATGCTGATCAATCACTGAGACCAACTTCAATCCCTTACATCTCTCTCTCTCTCTCTCT 1014  
 QY 996 GAGTGTCTGATGAGGCT 1048  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 1015 GGGCTCATGAACCT 1068

RESULT 9  
 AAL43942  
 ID AAL43942 standard; DNA; 1011 BP.  
 XX  
 AC AAL43942;  
 XX  
 DT 27-SEP-2002 (first entry)  
 XX  
 DE Human G protein-coupled receptor coding sequence.  
 XX  
 KW Human; gene therapy; G protein-coupled receptor; drug development;  
 KW central nervous system disease; endocrine disease; metabolic disease;  
 KW cancer; respiratory disease; digestive disease; immune disease;  
 KW inflammation; infection; circulatory disease; gene; ds.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 1..1011  
 FT /\*tag= a  
 FT /partial  
 FT /product= "Human G-protein coupled receptor"  
 FT /note= "No stop codon is given"  
 XX  
 PN WO200257441-A1.  
 XX



XX PS Claim 55; Page 113-114; 159pp; English.

XX CC The sequence encodes a human G-protein coupled receptor (GPCR),

CC hRUP21 The endogenous and non-endogenous, constitutively activated

CC versions of human G-protein coupled receptors (GPCR), are useful for

CC direct identification of candidate compounds as receptor agonists,

CC inverse agonists or partial agonists having applicability as therapeutic

CC agents for treating diseases related to GPCR, e.g. lung cancer.

CC Non-endogenous version of human GPCRs are also utilized in research

CC settings and in vitro and in vivo system, incorporating GPCRs can be

CC utilised to elucidate and understand the roles these receptors

CC play in the human condition, both normal and diseased.

XX SQ Sequence 1014 BP; 258 A; 263 C; 189 G; 304 T; 0 other;

Query Match 8.2%; Score 126.6; DB 22; Length 1014;

Best Local Similarity 49.9%; Pred. No. 6.7e-22;

Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

QY 60 CTTGTGAGAAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTT 119

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 59 CTTTGGAAATGCACTGATGAACATCCCACTCAAGATGCACTACCTCCCTGTATTT 118

QY 120 ATGCATCCAGTTCATTTTGGAGCTGCTTGGGAATGTCTACTGTGTGTGCGTACCTCT 179

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 119 ATGGCAATATCTCTCTGCTGGGATTCCAGGCAATGCAGTAGTATATCCACTTACATTT 178

QY 180 TCTGCATGAAGAACTGGAACACGACCAATGTCTATCTTTTAACTTTCCATCTCTGACT 239

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 179 TCAATGAGACCTTGGAGAGCAGCAGCATCATTTATGCTGAACCTGGCCTGCACAGATC 238

QY 240 TTGCTTTCCTGTCACCTTCCCATCTCTGATGAAGAGTATGCGCAAT---GATAAGGGGA 296

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 239 TGTGTATCTGACCAAGCTCCCTCTCTGATCTACTATGCTGCAAGTGGGAAACTGGA 298

QY 297 CCTATGGAGATGTTCTCTGTATGAACACCGATATGTCTTCAACCAACCTCTACACCA 356

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 299 TCTTTGGAGATTGATGTGATAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCA 358

QY 357 GCATCCTCTTCTCACTTCTATAGCATGGACGATATCTGCTCATGAAGTACCTTTCC 416

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 359 GCATCCTCTTCTCACTTCTGATCATCTTCCGCTACTGTGATCATTTCAACCAATGA 418

QY 417 GAGACACTTCTCAAAAGAGGAATTTGCCAFITTTAATCTGCTGCTGTCTGGGCT 476

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 419 GCTGCTTTTCCATTCACAAAACTCGATGTCAGTGTGATGCTGCTGTGTGGATCA 478

QY 477 TAGTGACCTTAGAAGTTCTACCATGCTCACTTTCATCAATTCGTCCCAAAAGAGG 536

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 479 TTTCACTGGTAGCTGATTCGATGACCTTCTTGATCATCATCAACCAAGGACCAACA 538

QY 537 GCAGTAATCGATGATATCAAGTTCTTGGAAACCTTGAACCAATCTCATTTACAGCC 596

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 539 GATCAGCTGTCTCGACCTCCACCATGCGG-----ATGACTCAATCTATTAAGTGGT 592

QY 597 TCTGCTCACTTGTGGGCTTCTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 656

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 593 ACAACCTGATTTGACTGCAACTACTTTCTGCTTCCCTTGGTGTATAGTACACTTTGCT 652

QY 657 AGATGGTAGTCTCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 653 ATACCAGATATCCACACTCTGACCATGAGCTGCAAACTGACAGTGGCTTTAAGCAGA 712

QY 717 AACCCCAACGCTGGTGGTCTGCGGTTGTGATCTTCTCTCTCTCTCTCTCTCTCTCTCT 776

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 713 AAGCACGAAGGCTAACCAATCTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 772

QY 777 ATATCATGCCAATTTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811

DB IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

DB 773 ATATCTTTGAGGGTATTCGGATCGAATCTCGCGTG 807

RESULT 11

ABN85630

ID ABN85630 standard; DNA; 1014 BP.

XX AC ABN85630;

XX DT 18-SEP-2002 (first entry)

XX DE Human P2Y-like receptor variant encoding gene SEQ ID NO 3.

XX KW Human; P2Y-like receptor; HIPHUM 0000037; immunity; inflammation;

XX KW cancer; Crohn's disease; irritable bowel syndrome; rheumatoid arthritis;

XX KW immunomodulator; anti-inflammatory; cytosolic; antitumorigenic;

XX KW gastrointestinal; anti-ulcer; antirheumatic; antithrombotic;

XX KW antibacterial; immunosuppressive; dermatological; nephrotoxic;

XX KW antiallergic; analgesic; receptor; gene; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..1014

XX FT /\*tag= a

XX FT /product= "P2Y-like receptor variant"

XX XX GB2369364-A.

XX PD 29-MAY-2002.

XX XX 31-AUG-2001; 2001GB-0021215.

XX PF 01-SEP-2000; 2000GB-0021524.

XX PR 06-SEP-2000; 2000GB-0021894.

XX PR 25-SEP-2000; 2000GB-0023444.

XX XX (GLAX ) GLAXO GROUP LTD.

XX XX Foord SM, Ignar DM;

XX XX WPI: 2002-511268/55.

XX XX P-PSDB; ABB83819.

An isolated P2Y-like receptor polypeptide (HIPHUM 0000037) which can be used for the identification of agonists and antagonists which may be used to treat an immune or inflammatory disease.

Claim 5; Page 28-29; 35pp; English.

The invention relates to an isolated P2Y-like receptor polypeptide (ABB83818-ABB83819) which is also referred to in the specification as HIPHUM 0000037. An effective amount of a substance (agonist or antagonist) which modulates P2Y receptor activity is useful to treat a subject having a disorder that is responsive to P2Y-like receptor modulation. The disorder is a disease of immunity or inflammation. The substance may also be used to manufacture a medicine for the treatment or prophylaxis of a disorder that is responsive to stimulation or modulation of P2Y-like receptor activity. Disorders which may be treated include colon cancers, asthma, COPD, Crohn's disease, irritable bowel syndrome, gastroenteritis and colitis, inflammatory bowel syndrome, ulcerative colitis, rheumatoid arthritis, viral diseases, bacterial infections, autoimmune diseases, dermatitis, glomerulonephritis allergies, allergic rhinitis, inflammatory pain and general inflammation such as tendonitis, polymyositis or prostatitis. The invention provides alternative substances for the treatment of immunological and inflammatory diseases. The present sequence is that the P2Y-like receptor variant encoding gene of the invention.

Query Match 8.2%; Score 126.6; DB 24; Length 1014;

Best Local Similarity 49.9%; Pred. No. 6.7e-22;

Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

QY 60 CTTGTGAGAAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTT 119







Db 539 GATCAGCCTGCTCGACCTCACCAGTTCGG-----ATGCACTCAATACTATTAAAGTGGT 592

QY 597 TCTGCCTGACTTGTGGCTTCCTAAATTCCTCTCTGTGATGCTCTTCTACTACA 656

Db 593 ACAACCTGATTTGACTGCACTACTTCTGCTCCCTTGTGTATAGTACACTTTGCT 652

QY 657 AGATGCTAGTCTTTAAAGAGGAGGAGCCAGCAGCACTGCCCTGCCACTGGACA 716

Db 653 ATACCAGGATATCCACACTCTGACCCATGAGCTGCAACTGACAGCTGCCCTTAAGCAGA 712

QY 717 ACCCCCAAGCCTGGTGGCTGGGGTGTGATCTTCTTACTTCTTCACTCTTCAACCCCTATC 776

Db 713 AAGCAGGAGGCTAACCATCTGCTACTCTTCTGCAATTTTACGTATGTTTTTACCCCTCC 772

QY 777 ATATCATGGCAATTTGAGGATCGCCTCACGCTG 811

Db 773 ATATCTTGAGGTCATTCGATCGAATCTCGCTG 807

RESULT 14

ABL56197

ID ABL56197 standard; cDNA; 1288 BP.

XX ABL56197;

AC.

DT 05-JUL-2002 (first entry)

XX Human P2Y1-11 encoding cDNA.

DE Human; P2Y1-11; chromosome 13; G protein-coupled; receptor;

KW gene therapy; thyroid; gene; ss.

KW

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT 18..1031

FT CDS

FT /\*tag= a

FT /product= "P2Y1-like purine receptor"

XX

PN DE10046970-Al.

XX

XX 11-APR-2002.

XX

XX 22-SEP-2000; 2000DE-1046970.

XX

XX 22-SEP-2000; 2000DE-1046970.

PR

XX (BRUE/) BRUESS M.

PA (BOEN/) BOENISCH H.

XX

XX Bruess M, Boenisch H;

PI

XX WPI; 2002-353329/39.

DR P-PSDB; ABB79438.

XX

XX New human P2Y111 gene, useful for treatment and diagnosis of associated diseases, and related proteins, antibodies and modulators, encodes G protein-coupled receptor

PT

XX Claim 5; Page 3; 5pp; German.

PS

XX The invention relates to the human P2Y111 gene (I), including its 5' and 3' untranslated regions, located on chromosome 13 and encoding a G protein-coupled receptor. (I) and related mRNA, cDNA, protein, antibodies etc., are used for diagnosis and (gene) therapy of diseases that are (indirectly) associated with (I) or its expression products. No diseases are specified but as (I) is expressed only in thyroid tissue, (I) is presumed to be involved in regulation of thyroid function. The present sequence is that of the P2Y111 encoding cDNA.

CC

XX Sequence 1288 BP; 347 A; 318 C; 241 G; 382 T; 0 other;

SQ

Query Match 8.2%; Score 126.6; DB 24; Length 1288;

Best Local Similarity 49.9%; Pred. No. 7.4e-22;

Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

QY 60 CTTGTGAGATTTGTTGGCAACAGAGCTATCTTGATTAAGTACTACCTCTCTGCATTTT 119

Db 76 CTTTGGAAATTTGCACTGATGAAACATCCCACTCAAGATGCACTACCTCCCTGTTATTT 135

QY 120 ATGCAATCGAGTTTCATTTTGGACTGCTTGGAAATGTCATCTGTTGTTTCGGTACTCTCT 179

Db 136 ATGCATTAATCTCTCTGCTGGGATTTCCAGGCAATGCACTAGTATGATATCCACTTACATTT 195

QY 180 TCTGCATGAAGAAGTGAACAGCAGCAATGCTATCTTTTAACTTTTCACTCTCTGACT 239

Db 196 TCAAAATGAGACCTTGAAGAGCAGCACCATTATGCTGAACCTGGCCTGCACAGATC 255

QY 240 TTGCTTTTCCGTGACCTTCCCATCTCTGATAAGAGTTTATGCCAAT---GATAAGGGGA 296

Db 256 TGCTGTATCTGACAGGCTCCCTTCTCTGATTTACTACTATGCGAGTGGGGAACATGGA 315

QY 297 CCTATGGAGATGTTCTCTGTATAAGCAACCGATATGCTTTCACACCACTCTTACACCA 356

Db 316 TCTTGGAGATTTTCATGTGTAAGTTTATCCGCTTCAGCTTCCCAATTTCAAGCTGTATAGA 375

QY 357 GCATCCTCTTCTCCTCATTATGATGAGCAGCATCTGCTCATGAAGTACCTTTTCC 416

Db 376 GCATCCTCTTCTCCTCATTATGATGAGCAGCATCTGCTCATGAAGTACCTTTTCC 435

QY 417 GAGAACACTTTCTACAAAGAGGAATTTGCCATTTTAAATCTCGCTGCTGTGGGCT 476

Db 436 GCTGCTTTTCCATTCACAAACTCGATGTCAGTTGTAGCTGTGCTGTGGTATCA 495

QY 477 TAGTGACCTTAGAAGTTCTACCAATGCTACTTTTCATCAATTTCTGCCAAAGAGAGG 536

Db 496 TTTCACTGGTAGCTGTCATTCGGATTTCTTGTATCATCAACCAAGGACCAACA 555

QY 537 GCAGTAAGTGCATGACTATGCAAGTTCTGGAACCTCTGAAACCTCAATCTCTATACAGCC 596

Db 556 GATCAGCCTGCTCGACCTCACCAGTTCGG-----ATGCACTCAATCTATTAAAGTGGT 609

QY 597 TCTGCTGACTTTTGTGGCTTCTCTAAATTCCTCTCTGCTGATGCTGCTTCTACTACA 656

Db 610 ACAACCTGATTTTGTACTGCACTACTTCTGCTCCCTTGGTGTATGATGACTTTGCT 669

QY 657 AGATGGTAGTCTTTAAAGAGGAGGAGCCAGCAGCACTGCCCTGCCACTGGACA 716

Db 670 ATACCAGGATTTATCCACACTCTGACCCATGACTGCAAACTGACAGCTGCTTAAAGAGA 729

QY 717 AACCCCAAGCCTGGTGGTCTGCGGTTGTGATCTTCTTATATCTTCTACACCTATC 776

Db 730 AAGCAGGAGGCTAACCAATTCCTGCTACTCTTGCATTTTACGTATGTTTTTACCCCTCC 789

QY 777 ATATCATGGCAATTTGAGGATCGCCTCACGCTG 811

Db 790 ATATCTTGAGGTCATTCGGATCGAATCTCGCTG 824

RESULT 15

AAS08362

ID AAS08362 standard; cDNA; 1729 BP.

XX AAS08362;

AC

XX 26-SEP-2001 (first entry)

DT Human cDNA encoding G-protein coupled receptor, GPCR 39404.

DE

XX Human; G-protein coupled receptor; GPCR; 39404; immunogen; antibody;

KW Good pasture's syndrome; viral hepatitis; Alzheimer's disease; lymphoma;

KW Heymann nephritis; Paget's disease; Crohn's disease; endometriosis;

KW systemic lupus erythematosus; actinic keratosis; myocarditis;

KW Kawasaki syndrome; DiGeorge syndrome; peripheral B-cell neoplasm;

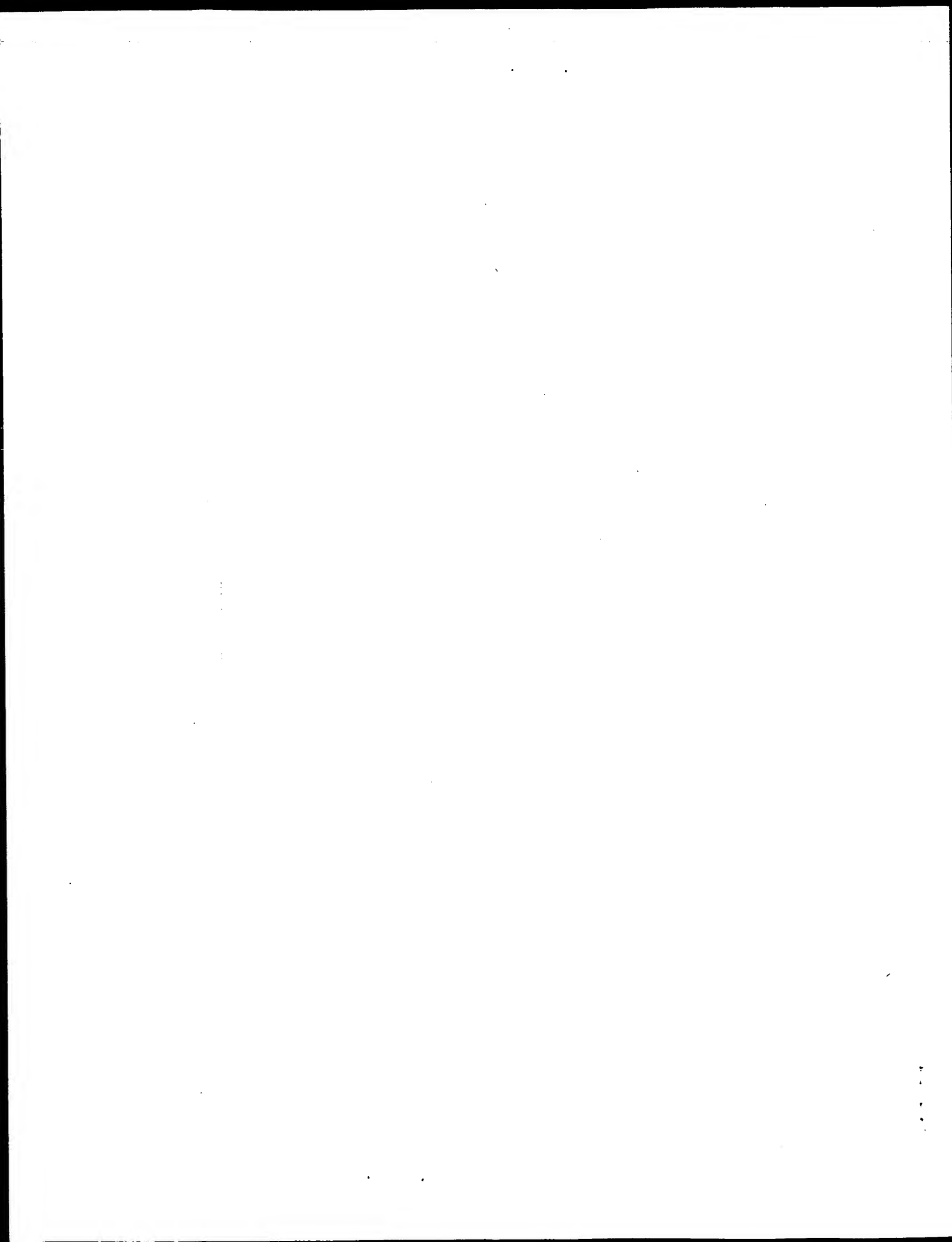
KW inflammations; teratoma; ss.

XX

OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH CDS 294..1307  
 FT /\*tag= a  
 FT /product= "Protein 39404"  
 XX  
 XX WO200149847-A2.  
 XX 12-JUL-2001.  
 XX  
 XX 22-DEC-2000; 2000WO-US35309.  
 XX  
 XX 30-DEC-1999; 99US-0475790.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Glucksmann MA, White D;  
 XX  
 XX WPI: 2001-432880/46.  
 XX P-PSDB; AAU04584.  
 XX  
 XX Novel isolated 26904, 38911 and 39404 polypeptides which are seven  
 PT transmembrane proteins belonging to superfamily of G-protein-coupled  
 PT receptors, useful for treating disorders of spleen, lung, liver, brain  
 PT and kidney -  
 XX  
 XX Claim 2; Fig 1; 164pp; English.  
 XX  
 XX The sequence encodes a novel human seven transmembrane domain  
 CC protein belonging to the G-protein coupled receptor (GPCR) superfamily,  
 CC protein 39404. The receptor is useful in drug screening assays, to  
 CC identify compounds that modulate receptor activity and/or interact with  
 CC the receptor, and for producing antibodies specific for the receptor, its  
 CC regions or fragments. The receptor is useful for treating/diagnosing a  
 CC 26904, 38911 and 39404 protein-associated disorder characterised by  
 CC aberrant expression or activity of the protein, for monitoring  
 CC therapeutic effect during clinical trials and other treatment, as bait  
 CC proteins in a two-hybrid or three-hybrid assay, and in pharmacogenomic  
 CC analysis. The proteins and nucleic acids encoding them are useful for  
 CC diagnosis and treatment of disorders selected from disorders of the  
 CC spleen, lung such as Good pasture's syndrome, liver such as viral  
 CC hepatitis, brain such as Alzheimer's disease, haematopoietic stem cells  
 CC such as lymphoma, kidney such as Heymann nephritis, bone such as Paget's  
 CC disease, colon such as Crohn's disease, uterus and endometrium such as  
 CC endometriosis, T-cell disorders such as systemic lupus erythematosus,  
 CC diseases of the skin such as actinic keratosis, disorders of the heart  
 CC such as myocarditis, disorders involving blood vessels such as Kawasaki  
 CC syndrome, disorders involving the thymus such as DiGeorge syndrome,  
 CC disorders involving B-cells such as peripheral B-cell neoplasms,  
 CC disorders of the breast such as inflammations, and disorders involving  
 CC the testis and epididymis such as teratoma. Numerous examples of  
 CC each type of disorder are given in the specification.  
 XX  
 XX Sequence 1729 BP; 461 A; 403 C; 302 G; 563 T; 0 other;  
 SQ  
 Query Match 8.2%; Score 126.6; DB 22; Length 1729;  
 Best Local Similarity 49.9%; Pred. No. 8.4e-22;  
 Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;  
 QY 60 CTTGTGAGATTGGTGGCAACAGAGGCTATCTTGAATAGTACTACCTCTCTGCAATTTT 119  
 DB 352 CTTTGGAAATTCGACTGATGAAGAACCCCTCAAGATGCACTACCTCCCTGTTATTT 411  
 QY 120 ATGCAATCGAGTTTCATTTTGGAGTCTGTTGGAAATGCTCACTGTGGGTTCGGCTACCTCT 179  
 DB 412 ATGGCATTTATCTCTCGTGGGATTTCCAGGCAATGCAGTAGTGATATCCACTTACATTT 471  
 QY 180 TCGCATGAAGAAATGGAACAGCAGCAATGCTCTATCTTTTAACTTTTCCATCTCTGACT 239  
 DB 472 TCAAAATGAGACCTTGGAAAGAGCAGCAGCATCATTTATGCTGAACCTGGCCTGCACAGATC 531  
 QY 240 TTGCTTTCTGTGCACCCCTCCCATCCTGATAAGAGATTATGCCAAT---GATAAGGGGA 296

Db 532 TGCTGTATCTGACCAAGCCTCCCTTCTCTGATTCACCTACTATGCCAGTGGCGAAACTGGA 591  
 QY 297 CCTATGGAGATGTTCTGTATATAAGCAACCGATATGCTTACACCAACCTCTTACACCA 356  
 Db 592 TCTTTGGAGATTTTCATGTGTAAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCA 651  
 QY 357 GCATCTCTTCCCTCACTTTTCATTAGCATGGACGGATATCTGCTCATGAAGTACCCCTTCC 416  
 Db 652 GCATCTCTTCCCTCACTTTTCAGCATCTTCCGCTACTGTGTGATCATTCACCAATGA 711  
 QY 417 GAGAACACTTTTCTACAAAAGAGATTTGCCATTTTAAATCTCGCTGGCTGTCTGGGCT 476  
 Db 712 GCTGCTTTTCCATTCACAAAACCTCGATGTGAGTTTCTAGCCTGTGCTGGGTGTGATCA 771  
 QY 477 TAGTGACCTTAGAAGTTTCTACCCATGCTCACTTTTCATCAATTTCTGCTCCAAAAGAGAGG 536  
 Db 772 TTTCACTGGTGTGCTCATTTCCGATGACCTTCTTGATCACATCAACCAACAGGACCAACA 831  
 QY 537 GCAATAACTGCAATCGACTATGCAAGTTCTGGAACCCCTGAACACAATCTCATTTACAGCC 596  
 Db 832 GATCACCTGCTCGACCTCACAGTTCCG-----ATGAACCTCAATACTATTAAAGTGT 885  
 QY 597 TCTGCTGACTTTGTTGGCTTCTTAATTCCTCTCTGTGTGATGTGCTTTCTTCTACTACA 656  
 Db 886 ACAACCTGATTTTGACTGCAACTACTTTCTGCTCCCTTGGTGTGATGTGACACTTTGCT 945  
 QY 657 AGATGCTAGTCTTCTTAAGAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACA 716  
 Db 946 ATACCACGATTTATCCACACTCTGACCCATGGSACTGCAAACTGACAGCTGCCCTTAAGCAGA 1005  
 QY 717 AACCCCAACGCTGTGTGGTGGCTGTGCTGCTCTCTCTATCTCTTACACCTATC 776  
 Db 1006 AAGCAGGAAGGCTAACCACTTCTGCTACTCTTGCATTTTACGTATGTTTTTACCCCTCC 1065  
 QY 777 ATATCATGCGCAATTTGAGGATCGCCTCAGCCTG 811  
 Db 1066 ATATCTTGAGGCTCATTCGGATCGAATCTCGCCTG 1100

Search completed: April 2, 2003, 13:08:37  
 Job time : 399 secs











Db 830 GCTGTTCTCCATGCGGTGATGCGCTTAAAGGTGAGGACGATCAGGATGGGCACAAACGCTG 889  
QY 458 TCCTGCTGCTCTGGGCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAAT 517  
Db 890 TGCTGGCAGTATGGCTAAACCGCATATATGCTACCATCCCATGCTAGTGTGTTTACCAA 949  
QY 518 TCTGTCCAAAAGAGGAGGCGAGTAACTGCATCGACTATGCAAGTCTTGGAAACCCCTGAA 577  
Db 950 GTGGCTCTGAAGATGCTGTTCTACAGTGTATTTCATTTTACAATCAACAGACTTTGAAG 1009  
QY 578 CACAATCTCATTTACAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 637  
Db 1010 TGAAGATCTTCACCAAACTTCAAAATGAACATTTTGGCTGTGTTGATCCCATTCACCATC 1069  
QY 638 ATGTGCTTCTTCTACTACAAGATGTTAGTCTTCTTAAAGAGGAG 681  
Db 1070 TTTATGTTCTGCTACATTAATCTCTGCACGAGCTGAAGAGTG 1113

RESULT 5  
US-08-513-974B-57  
; Sequence 57, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/513.974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236356

US-08-461-244-1  
; Sequence 1, Application US/08461244  
; Patent No. 5776729  
; GENERAL INFORMATION:  
; APPLICANT: Soppet, Daniel R.  
; APPLICANT: Yi, Li  
; APPLICANT: Ruben, Steven M.  
; APPLICANT: Rosen, Craig A.  
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
; ADDRESS: STUART & OLSTEIN  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461.244  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferraro, Gregory D.  
; REGISTRATION NUMBER: 36,134  
; REFERENCE/DOCKET NUMBER: 325800-445  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1586 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 431..1495  
US-08-461-244-1

Query Match  
Best Local Similarity 47.3%; Pred. No. 9.2e-13;  
Matches 276; Conservative 0; Mismatches 305; Indels 3; Gaps 1;

QY 98 AAGTACTACCTCTCGCATTTTATGCAATCGAGTTCTTTTGGAGCTGCTGGGAATGTC 157  
Db 533 AAGTTGCTCTGCTGCTGCTTTTATGCTCTCTGTTTCTATTCAGTCTCTGGGAACAGC 592  
QY 158 ACTGTGTGTGTTGGCTACCTCTCTGCTGAAGAACTCTGCAACAGCAATGCTATCTT 217  
Db 593 CTGGTCATCTCCCTGGTCTGTTGGTCTGCAAGAAGCTGAGGAGCATCACAGATGATACCTC 652  
QY 218 TTTAACTTTCCATCTCTGACTTTGCTTTCTGTGACCCCTTCCCACTCTGATAAAGAGT 277  
Db 653 TTGAACCTGGCCCTGCTGACCTCTCTTCTCTCTCTCCCTTTCAGACCTA---C 709  
QY 278 TATGCAATGATAAGGGACCTATGAGATGTTCTCTGTATAGCAACCGATATGTGCTT 337  
Db 710 TATCTGTGGACAGTGGGCTTTGGGACGCTATATGTGCAAGGCTGCTGGCTTTTAT 769  
QY 338 CACACCAACCTCTACACAGCATCCTCTCTCTCACTTTTCAATAGCATGGACCGATATCTG 397  
Db 770 TACATTGGCTTCTACAGAGCATGTTTTTTCATCACCCCTCATGAGTGTGGACAGGTACCTG 829  
QY 398 CTCATGAAGTACCCCTTTCCGAGACACTTTCTACAAAAGAAGAAATTTGCCATTTTAATC 457

QY	737	CTGGCGTTGTGATCTTCTATACCTTTACACACCCATATCATATCATGCGCAATTTGAGG	799
Db	727	GTGGTGGCTGCTGCTGCTTTGCCATCAGCTTCTCTGCTTTTTCACATCACCACGACCTAC	786
QY	797	ATGCGCTCACGCGCTGATAGTTGGCCACAGGATGTACACAGAGCCCATCAATCTATA	856
Db	787	CTGGCAGTGGCTGACGCGCGGCGCTCCCTGCATGTATTGGAGGCTTTTGCAGCGGCC	846
QY	857	TACACACTGACGCGGCTCTGGCCCTTTCTGAACAGTGGCCATCAATCCCATCTTCTACTTC	916
Db	847	TACAAAGGACGCGCGCTTTGCCAGTGCCAAACAGCGTCTGGAGCCCATCTCTCTTCTAC	906
QY	917	CTCA 920	
Db	907	TTCA 910	
<p>RESULT 6</p> <p>US-08-513-974B-379</p> <p>; Sequence 379, Application US/08513974B</p> <p>; Patent No. 6114139</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Hinuma, Shuji</p> <p>; APPLICANT: Hosoya, Masaki</p> <p>; APPLICANT: Fujii, Ryo</p> <p>; APPLICANT: Ohtaki, Tetsuya</p> <p>; APPLICANT: Fukusumi, Shoji</p> <p>; APPLICANT: Ohgi, Kazuhiro</p> <p>; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,</p> <p>; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF</p> <p>; NUMBER OF SEQUENCES: 380</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP</p> <p>; STREET: 130 Water Street</p> <p>; CITY: Boston</p> <p>; STATE: MA</p> <p>; COUNTRY: USA</p> <p>; ZIP: 02109</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/513,974B</p> <p>; FILING DATE: 14-SEP-1995</p> <p>; CLASSIFICATION: 536</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: PCT/JP95/01599</p> <p>; FILING DATE: 10-AUG-1995</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 7-093989</p> <p>; FILING DATE: 19-AUG-1995</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 7-057186</p> <p>; FILING DATE: 16-MAR-1995</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 7-007177</p> <p>; FILING DATE: 20-JAN-1995</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 6-326611</p> <p>; FILING DATE: 28-DEC-1994</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 6-270017</p> <p>; FILING DATE: 02-NOV-1994</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 6-236357</p> <p>; FILING DATE: 30-SEP-1994</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: JP 6-236356</p> <p>; FILING DATE: 30-SEP-1994</p> <p>; PRIOR APPLICATION DATA:</p>			

Db 703 GATGCCCGGCGAGAGCCTGTGGCCCGAGGAGCGGGGTGGCAAGCGCGCCCGCATGGCCGTC 762

QY 737 CTGGCGTTGTGATCTTCTCTATACCTTTTACACACCCCTATCATATCATGCGCAATTGAGG 796

Db 763 GTGGTGGCTGCTGCTTGGCCATCAGCTTCTGCTTTTTCACATCACCAGACAGCCTAC 822

QY 797 ATGCGCTTACGCGCTGGATAGTTGGCCACAGGATGTACACAGAGCCCATCAATCTATA 856

Db 823 CTGGCAGTGGGCTCGAGCGCGGGCGTCCCTGCACCTGTATTGGAGCCCTTTCGAGCGGCC 882

QY 857 TACACACTGACAGCGCTCTGCGCTTTCGACAGTGGCCATCAATCCCATCTTCTACTTC 916

Db 883 TACAAGGCGACGCGGGCGGTTTGGCAGTGGCCACAGCGGTGCTGACGCCCATCTCTTCTAC 942

QY 917 CTCA 920

Db 943 TTCA 946

RESULT 7

US-08-012-988A-1

; Sequence 1, Application US/08012988A

; Patent No. 5652133

; GENERAL INFORMATION:

; APPLICANT: Murphy, Phillip M.

; TITLE OF INVENTION: Cloning and Expression of Human

; TITLE OF INVENTION: Macrophage Inflammatory Protein-1 alpha (MIP-1

; TITLE OF INVENTION: alpha)/RANTES Receptor

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Kourie and Crew

; STREET: One Market Plaza, Steuart Tower, Suite 2000

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94610

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/012,988A

; FILING DATE: 19930128

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Kenneth A.

; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 15280-118

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-543-9600

; TELEFAX: 415-543-5043

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2156 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: primer\_bind

; LOCATION: 259..275

; FEATURE:

; NAME/KEY: primer\_bind

; LOCATION: complement (868..884)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 63..1128

; US-08-012-988A-1

Query Match 5.1%; Score 78.4; DB 1; Length 2156;

Best Local Similarity 54.1%; Pred. No. 3,1e-12;

Matches 160; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 89 ATCTTGAATGAAGTACTACCTCTCTGATTTTATGCAATCGAGTTCAATTTTGGACTGCTT 148

Db 103 AACTTCAAGCAACTGTGCTGCCACCTGTGATTCGGCGGTGCTGCGGCTGCGCTGCG 162

QY 149 GGAATGCTACTGTGTGCTGCTACCTCTTCTGATGAAGAACTTGAACAGCAGCAAT 208

Db 163 CTGACATCTGTGCTATACCCAGATCTGCACGTCGCGCGGCCCTGACCGCGCAGCGCC 222

QY 209 GTCTATCTTTTAACTTTCATCTCTGACTTTCCTTCTGACACCTTCCCTCCATCTG 268

Db 223 GTGTACACCTAACTTGTCTGGCTGACCTGCTATATGCTGCTCTCCCTGCGCTGCTC 282

QY 269 ATAAGAGTTATGCAAA---TGATAAGGGACCTATGGAGATGTTCTGTATATAAGCAAC 325

Db 283 ATCTACAACTATGCCAAGTGATCACTGCGCCCTTTGGGACCTTGGCTGCGCGCTGGTC 342

QY 326 CGATATGTGTTACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCT 385

Db 343 CGCTTCTCTCTATGCAACCTGCGAGGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 402

QY 386 GACGATATCTGCTCATGAAGTACCTTT---CCGAGAAGCACTTTCTACAAAAGAGAA 412

Db 403 CAGGCTTACCTGGGCTATCTGCCACCCCTGCGCCCTGGCCACAACTGGCGGGCGCCGG 462

QY 443 TTTGCCATTTTAACTCGCTGCTGCTGCGGCTTATGACCTTTAGTACCTTAGAAGTTTACCCATG 502

Db 463 GCTGCTGCTGAGTGTGTGAACGCTGTGGCTGGCGGTGACAAACCCAGTGCCTGCCACA 522

QY 503 CTCAGTTTCATCAATTTGTTCCCAAGAGAGGAGGCGGTAAGTGCATGCACTATGCAAGT 562

Db 523 GCCATCTTCTGCGCACAGGATCCAGGTAACCGGCACTGTCTGCTATGACCTACGCCCG 582

QY 563 TCTGGAACCTTGACACAACTCTCAATTTACAGCCTCTGCTGCTGCTTGTGTTGGCTTCTCTA 622

Db 583 CTGCGCTGGCCACCACTATATGCCATATGCGCTATGCGATGCTCTCACTGCTATCGGCTTCTCTG 642

QY 623 ATTCCTCTCTCTGATGCTGCTTCTTCTACTACAAGATGTTAGTCTCTTAAAGAGGAGG 682

Db 643 CTGCGCTTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702

QY 683 AGCCAGCAGCAAGCAACTG-----CCCTGCCACTGGGACAAACCCCAACGCGCTGGTGTCT 736

Query Match 5.2%; Score 79.6; DB 3; Length 1023;

Best Local Similarity 46.0%; Pred. No. 9.3e-13;

Matches 388; Conservative 0; Mismatches 444; Indels 12; Gaps 3;

QY 89 ATCTTGAATGAAGTACTACCTCTCTGATTTTATGCAATCGAGTTCAATTTTGGACTGCTT 148

Db 103 AACTTCAAGCAACTGTGCTGCCACCTGTGATTCGGCGGTGCTGCGGCTGCGCTGCG 162

QY 149 GGAATGCTACTGTGTGCTGCTACCTCTTCTGATGAAGAACTTGAACAGCAGCAAT 208

Db 163 CTGACATCTGTGCTATACCCAGATCTGCACGTCGCGCGGCCCTGACCGCGCAGCGCC 222

QY 209 GTCTATCTTTTAACTTTCATCTCTGACTTTCCTTCTGACACCTTCCCTCCATCTG 268

Db 223 GTGTACACCTAACTTGTCTGGCTGACCTGCTATATGCTGCTCTCCCTGCGCTGCTC 282

QY 269 ATAAGAGTTATGCAAA---TGATAAGGGACCTATGGAGATGTTCTGTATATAAGCAAC 325

Db 283 ATCTACAACTATGCCAAGTGATCACTGCGCCCTTTGGGACCTTGGCTGCGCGCTGGTC 342

QY 326 CGATATGTGTTACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCT 385

Db 343 CGCTTCTCTCTATGCAACCTGCGAGGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 402

QY 386 GACGATATCTGCTCATGAAGTACCTTT---CCGAGAAGCACTTTCTACAAAAGAGAA 412

Db 403 CAGGCTTACCTGGGCTATCTGCCACCCCTGCGCCCTGGCCACAACTGGCGGGCGCCGG 462

QY 443 TTTGCCATTTTAACTCGCTGCTGCTGCGGCTTATGACCTTTAGTACCTTAGAAGTTTACCCATG 502

Db 463 GCTGCTGCTGAGTGTGTGAACGCTGTGGCTGGCGGTGACAAACCCAGTGCCTGCCACA 522

QY 503 CTCAGTTTCATCAATTTGTTCCCAAGAGAGGAGGCGGTAAGTGCATGCACTATGCAAGT 562

Db 523 GCCATCTTCTGCGCACAGGATCCAGGTAACCGGCACTGTCTGCTATGACCTACGCCCG 582

QY 563 TCTGGAACCTTGACACAACTCTCAATTTACAGCCTCTGCTGCTGCTTGTGTTGGCTTCTCTA 622

Db 583 CTGCGCTGGCCACCACTATATGCCATATGCGCTATGCGATGCTCTCACTGCTATCGGCTTCTCTG 642

QY 623 ATTCCTCTCTCTGATGCTGCTTCTTCTACTACAAGATGTTAGTCTCTTAAAGAGGAGG 682

Db 643 CTGCGCTTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702

QY 683 AGCCAGCAGCAAGCAACTG-----CCCTGCCACTGGGACAAACCCCAACGCGCTGGTGTCT 736





Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:51 11 November 2014





INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 34..1116  
PCT-US93-09636-3

Query Match 4.8%; Score 74; DB 5; Length 1643;  
Best Local Similarity 47.4%; Pred. No. 4.6e-11;  
Matches 255; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

```
QY 126 TCGAGTTCATTTTGGAGTCTGGGAATGTCACCTGGTGGTTCGGCTACTCTTTCGCA 185
DB 152 TCGTTCATCATTTGGGCTCTGGGAACCTTACTAGCCTTGGTGGTCAATGTTCAAAACA 211
QY 186 TGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTCTTTCATCTCTGACTTTGCTT 245
DB 212 GGAAGAAATCACTACACCTCTATTCAACAAATTTGGTGATTTCTTGATATCTTT 271
QY 246 TCCTGTCCACCTTCCCATCTGATAAGAGTTATGCCAATG---ATAAGGGACCTATG 302
DB 272 TTACCACGGCTTTGGCTTACACGAATAGCCCTACTATGCAATGGCTTTGACTGGGAATCG 331
QY 303 GAGATGTTCTCTGTATAAGCAACCGATATGTCTTTCACACCAACCTCTACACCAGCATCC 362
DB 332 GAGATGCTTGTAGGATAACTGCGCTAGTGTATACATCAACATATGACGGTGTGA 391
QY 363 TCCTCTCACTTTCACTAGCATGGACGATATCTGCTCAATGAAGTACCTTTCCGAGAAC 422
DB 392 ACTTTATGACCTGCTGAGTATTGACCGCTTCACTGCTGGTGGACCTCTAGCGTACA 451
QY 423 ACTTTCTACAAAAGAGGAAATTTGCCATTTTATCTCGTGGCTGCTGGGCTTAGTGA 482
DB 452 ACAAGATAAAGAGATTGAACATGCAAAAGCGCTGTCATATTGCTGGATTTCTAGTAT 511
QY 483 CCTTAGAAGTTTACCCTACCTCTCATCAATTTCTGTCCTCCAAAGAGAGGCGAGTA 542
DB 512 TTGCTCAGACACTCCCACTCTCATCAACCTATGTCAAGCAGGAGGCTGAAGGATTA 571
QY 543 ACTGCATCCACTATGAAGTCTGGAACCCCTGGAACAAATCTCATTTACAGCCCTCTGCC 602
DB 572 CATGCATGGAGTATCCAAACTTTGAAGAACTAAATCTCTTCCCTGGATTTCTGCTGGG 631
QY 603 TGACTTTGTGGCTTCCTAAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 660
DB 632 CATGTTTCATAGGATATGTACTTCCACTTATATCATTTCTCATCTGCTATTCTCAGAT 689
```

## RESULT 15

US-08-847-296B-2  
Sequence 2, Application US/08847296B  
Patent No. 6271347

## GENERAL INFORMATION:

APPLICANT: DAUGHERTY, BRUCE L.  
APPLICANT: DEMARTINO, JULIE A.  
APPLICANT: SICILIANO, SALVATORE J.  
APPLICANT: SPRINGER, MARTIN J.  
TITLE OF INVENTION: EOSINOPHIL EOTAXIN RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/847,296B  
FILING DATE: 24-APR-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/016,158  
FILING DATE: 26-APR-1996

APPLICATION NUMBER: 60/017,113  
FILING DATE: 26-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Thies, J. Eric

REGISTRATION NUMBER: 35,382

REFERENCE/DOCKET NUMBER: 19634Y

TELEPHONE: 908-594-3904

TELEFAX: 908-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1065 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-847-296B-2

Query Match

Best Local Similarity 4.8%; Score 73.8; DB 4; Length 1065;

Mismatches 198; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 117 TTATGCAATCGAGTTTCATTTTGGAGTCTTGGGAATGTCACCTGGTGGTTCGGCTACC 176

DB 119 TGTATCTCCCTGGTGTGTCTCACTGTGGCCCTCTTGGCAATGTGGTGGTGTATGATCTCA 178

QY 177 TCTTCTGCATGAAGAACTGGAAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTG 236

DB 179 TAAATACAGGAGGCTCGGAATTAAGCAACATCTACCTCTCACTGCGCAATTCGG 238

QY 237 ACTTGTCTTCTGTCGACCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGGA 296

DB 239 ACTGTCTCTCTCTGCTCACCTTCCATCTGGATCCCACTATGTCAGGGGGCATAAATCTGG 298

QY 297 CCTATGAGAGTGTCTCTGTATAGCAGCAGCAATGTCTTCAACCAACCTCTACACCA 356

DB 299 TTTTGGCCATGGCATGTGAAGCTCTCTCAGGGTTTATACACAGGCTTGTACAGCG 358

QY 357 GCATCTCTTCTCTCACTTTCAATAGCAGTGGACCGATATCTGCTCATGAAGTACCTTTCC 416

DB 359 AGATCTTTTTCATAATCTCTGCTGACAAATCGACAGTACCTGGCCATTTGCCATCTGTGT 418

QY 417 GAGAACACTTTTACAAAGAGGAATTTGCCATTTTAACTCTGCTGGTGTCTGGGCGCT 476

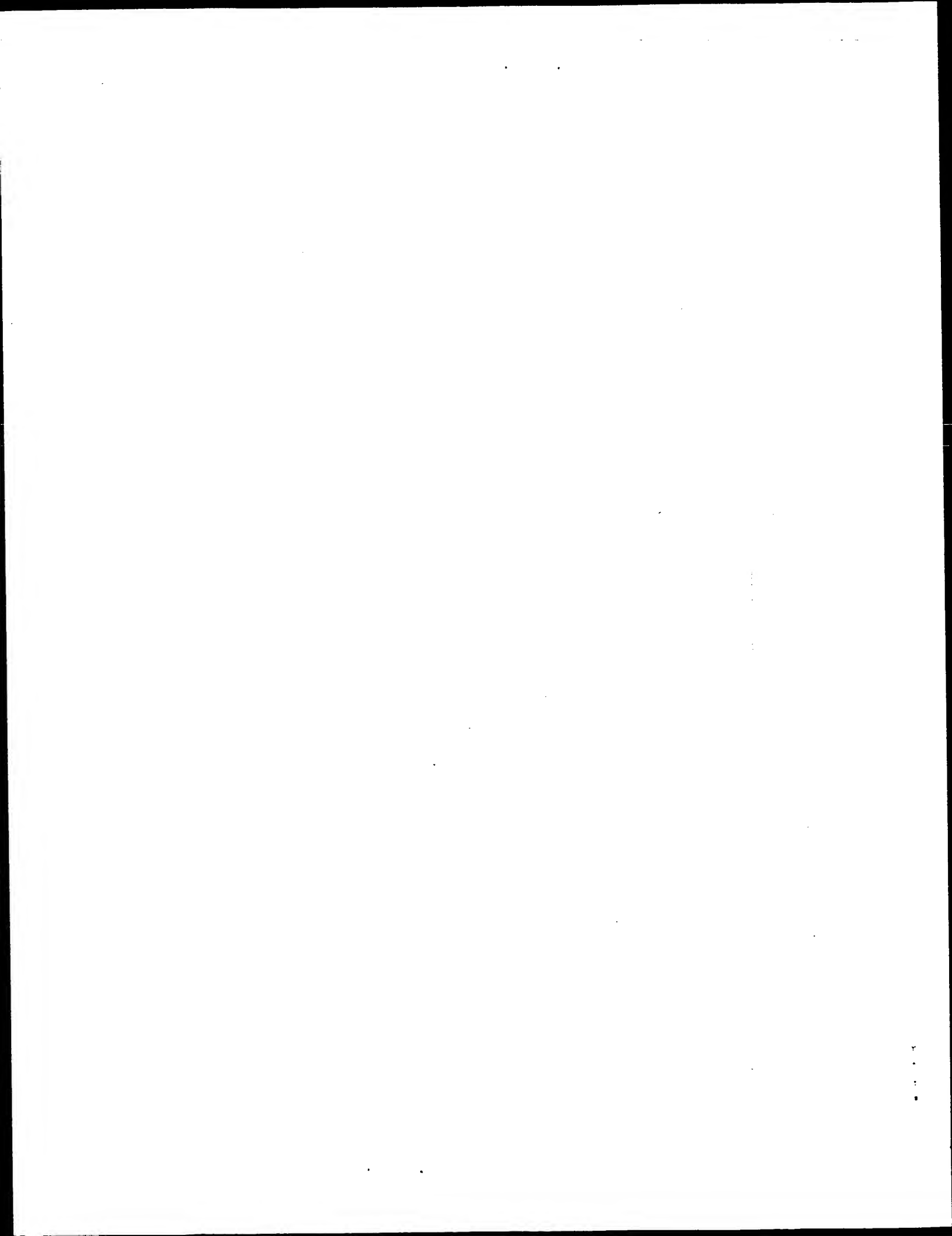
DB 419 TTGCCCTTCGAGCCCGGACTGTCTACTTTTGTGTGTATCATCAGCAGCATCTCACCTGGGCG 478

QY 477 TAGTGACCTTAGAAGTTTACCCCATGCTCACTTTTCATCAATTCG 521

DB 479 TGGCAGCTAGCAGCTCTTCTCTGAATTTATCTTCTATGAGACTG 523

Search completed: April 2, 2003, 15:13:46

Job time : 79 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2003, 13:02:01 ; Search time 154 Seconds  
(without alignments)  
8788.765 Million cell updates/sec

Title: US-09-891-138A-1  
Perfect score: 1543  
Sequence: 1 gctctggcagagtgttctg.....tgctaaataaatacaata 1543

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues  
Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications\_NA:  
1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq:  
12: /cgn2\_6/ptodata/2/pubna/US10\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	590.8	38.3	1428	9	US-10-270-587-1
2	126.6	8.2	1014	9	US-10-270-144-1
3	126.6	8.2	1014	10	US-09-943-798-3
4	126.6	8.2	9905	9	US-10-270-144-3
5	125	8.1	1014	9	US-10-023-775B-1
6	119	7.7	1313	10	US-09-728-422-1
7	96	6.2	831	10	US-09-943-798-1
8	82.2	5.3	1805	10	US-09-823-114-18
9	82.2	5.3	1829	9	US-09-905-186A-9
10	82.2	5.3	1829	9	US-09-905-186A-10
11	82.2	5.3	2534	9	US-09-905-186A-11
12	82.2	5.3	2534	9	US-10-087-345A-22
13	80.6	5.2	1829	9	US-09-905-186A-7
14	80.6	5.2	1829	9	US-09-905-186A-8
15	80	5.2	1586	10	US-09-104-792-1
16	77.6	5.0	1773	9	US-10-112-599A-3
17	76.4	5.0	2602	9	US-09-905-186A-1
18	74.2	4.8	1301	9	US-10-024-494-7
19	73.8	4.8	1065	9	US-09-922-895-2

20	73.8	4.8	1689	10	US-09-931-381A-15
21	73.8	4.8	1717	10	US-09-964-824A-100
22	73.8	4.8	1915	12	US-10-106-623-3
23	73.8	4.8	3426	9	US-10-001-835-29
24	71.2	4.6	2218	10	US-09-214-904-3
25	71.2	4.6	2219	9	US-10-112-599A-1
26	71	4.6	1427	10	US-09-967-768A-296
27	70.8	4.6	1607	9	US-10-120-394-19
28	70.8	4.6	1607	9	US-09-764-413-19
29	70.8	4.6	1677	10	US-09-837-446-1
30	69.4	4.5	1895	9	US-09-940-240-19
31	69.2	4.5	1860	9	US-09-940-240-17
32	69	4.5	1423	9	US-10-116-252-3
33	69	4.5	1594	10	US-09-826-508-39
34	68.4	4.4	2080	12	US-10-044-090-628
35	68.4	4.4	2081	12	US-10-044-090-627
36	68	4.4	1829	10	US-09-823-114-7
37	67.8	4.4	1318	10	US-09-917-800A-1445
38	66.6	4.3	1487	10	US-09-789-482-3
39	66.6	4.3	1487	10	US-09-789-486-3
40	66.6	4.3	3100	10	US-09-954-456-267
41	66.6	4.3	3100	10	US-09-954-456-945
42	66.6	4.3	3100	10	US-09-954-456-1588
43	66.4	4.3	526	10	US-09-812-102-40
44	65.2	4.2	2050	9	US-09-940-240-15
45	63.8	4.1	1143	10	US-09-997-522-1

## ALIGNMENTS

### RESULT 1

US-10-270-587-1  
; Sequence 1, Application US/10270587  
; Publication No. US20030054487A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Yi  
; TITLE OF INVENTION: Human G-Protein Coupled Receptor  
; FILE REFERENCE: PF217C2  
; CURRENT APPLICATION NUMBER: US/10/270,587  
; CURRENT FILING DATE: 2002-10-16  
; PRIOR APPLICATION NUMBER: US 09/908,593  
; PRIOR FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 08/781,456  
; PRIOR FILING DATE: 1997-01-10  
; PRIOR APPLICATION NUMBER: US 60/009,902  
; PRIOR FILING DATE: 1996-01-11  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1428  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-270-587-1

Query Match 38.3%; Score 590.8; DB 9; Length 1428;  
Best Local Similarity 75.2%; Pred. No. 2.5e-135;  
Matches 763; Conservative 0; Mismatches 247; Indels 4; Gaps 2;

QY	39	CCAGATGGCAGACAAATTATCTTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATA	98
DB	99	GGATCATGGCATGGAATGCAACTTGCACAACTGGCTGGCAGCAGAGGCTGCCTGGAAA	158
QY	99	AGTACTACTCTCTGCAATTTATGCAATCGAGTTCATTTTGGACTGCTTGGGAATGTCA	158
DB	159	AGTACTACTCTTCCATTTTATGGGATTCAGTTCGTTGGGAGCTCTTGGAAATATCCA	218
QY	159	CTGTGGTGTTCGGCTACTCTCTTCGATGAAGAACTGGACACAGCAATGTCTATCTTT	218
DB	219	TGTGTGTGTACGGCTACATCTTCTCTCTGAAGAACTGGACACAGCAATGTCTATCT	278
QY	219	TTAACCCTTCATCTCTGAGCTTTGCTTTCTGTGCACCCCTGCCATCTGATAAAGAGTT	278

```

Db 279 TTAACCTCTCTGCTGCTGACTTAGCTTTTCTGTGACACCTCCCATGCTGATAGAGATT 338
QY 279 ATGCCAATGATAGGGAGCTATGAGAGATGTTCTCTGTATAGCAACCGATATGTCCTTC 338
Db 339 ATGCCAATGGAACCTGATATATGAGAGCTGCTCTGCAATAGCAACCGATATGTCCTTC 398
QY 339 ACACCACTCTACACGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
Db 399 ATGCCAACCTCTATACAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 458
QY 399 TCATGAAGTACCTCTCCGAGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 458
Db 459 TAATTAAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 518
QY 459 CGCTGGCTGTCTGGGCTTTAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 518
Db 519 CTTTGGCCATGTGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 578
QY 519 CTGTCCCAAGAGAGGAGGAGTAACTGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 578
Db 579 CTGTATTAACCTGCAATGGCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 638
QY 579 ACAATCTCAATTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 638
Db 639 ACAACCTCAATTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 698
QY 639 TGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 698
Db 699 TGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 758
QY 699 CTGCTCTGCACTGGACAAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 758
Db 759 CTGCTCTGCACTGGACAAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 818
QY 759 TACTCTTTCACACCTATCATATCATGCGCAATTTGAGGATCGCTCTCTCTCTCTCTCTCT 818
Db 819 TGTCTTTTACACCTATCATGCTGATGCGCAATTTGAGGATCGCTCTCTCTCTCTCTCTCT 878
QY 819 G---GCCAAGGATGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875
Db 879 GGAAGCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 938
QY 876 TGGCTTTCTGACAGCTGCCATCAATCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 935
Db 939 TGGCTTTCTGACAGCTGCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 998
QY 936 GAGAGATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 995
Db 999 GGGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1058
QY 996 GAGCTGCTGGATGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1048
Db 1059 GGGCTCATGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1112

```

```

RESULT 2
US-10-270-144-1
; Sequence 1, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1014

```

```

; TYPE: DNA
; ORGANSIM: Human
US-10-270-144-1
Query Match      8.2%; Score 126.6; DB 9; Length 1014:
Best Local Similarity 49.9%; Pred. No. 2.9e-21;
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps
QY 60 CTCTGAGAAATGTTGGCAACAGAGGCTATCTTTGAATAAGTACTACTCTCTCTCTCTCTCT 119
Db 59 CTTTGGAAATGCTGACTGATAAATCCACCTCAAGATGCACTACTCTCTCTCTCTCTCTCT 118
QY 120 ATGCAATCGAGTTTCATTTTGGGACTGCTTTGGGAATGTCAGTGTGTTGGTCTCTCTCTCT 179
Db 119 ATGCGATTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 178
QY 180 TCTGCATGAAGAACTGGAACAGCAGCAATGCTATCTTTTAACTCTTCTCTCTCTCTCTCT 239
Db 179 TCAAAATGAGACCTTTGGAAAGAGCAGCACCATTATGCTGCAACCTGGCTGACAGATC 238
QY 240 TTGCTTTTCTCTGTCACCCCTTCCCATCTCTGATAAAGAGTTATGCAAT---GATAGGGGA 296
Db 239 TGCCTGATCTGACGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 298
QY 297 CTTATGAGAGATGTTCTCTGTATAGCAACCGATATGCTTTCACACCAACCTCTACACCA 356
Db 299 TCTTTGGAGATTTCAAGTGTAAAGTTTATCCCGCTTCAAGCTTCCAGCTTCCATCTGATAGA 358
QY 357 GCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 416
Db 359 GCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 418
QY 417 GAGAACACTTTTACAAAAGAGGAATTTGCCATTTTAACTCTCTCTCTCTCTCTCTCTCTCT 476
Db 419 GCTGCTTTTCTCAATTCACAAAACCTGATGTCAGTTTGTAGCCTGTGCTGTGGTGATCA 478
QY 477 TAGTGACCTTAGAAGTTCTTACCCATGCTTCACTTTTCACTCAATTTCTGTCCTCAAAAGAG 536
Db 479 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
QY 537 GCAGTAACCTGCACTGATGCAAGTTTCTGGAAACCTTGAACACAACTCTCTCTCTCTCTCT 596
Db 539 GATCAGCTGCTGCTGCACTGATGCAAGTTTCTGGAAACCTTGAACACAACTCTCTCTCTCT 592
QY 597 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
Db 593 ACAACCTGATTTTGAAGTCTGCAACTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 652
QY 657 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
Db 653 ATACCACGATTTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712
QY 717 AACCCCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
Db 713 AAGCAGGAAGGCTAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
QY 777 ATATCATGGCAATTTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811
Db 773 ATATCTTGGGGTCAATTCGATCGAATCTCGCGCTG 807

```

```

RESULT 3
US-09-943-798-3
; Sequence 3, Application US/09943798
; Patent No. US20020065215A1
; GENERAL INFORMATION:
; APPLICANT: Glaxo Group Limited
; TITLE OF INVENTION: Polypeptide
; FILE REFERENCE: QG1021
; CURRENT APPLICATION NUMBER: US/09/943,798
; CURRENT FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 3  
; LENGTH: 1014  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-943-798-3

Query Match 8.2%; Score 126.6; DB 10; Length 1014;  
Best Local Similarity 49.9%; Pred. No. 2.9e-21;  
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

```

QY 60 CTTGTGAGATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCAATTT 119
Db 59 CTTTGGAAATGGCAATGATGAATAATCCCACTCAAGATGCACCTACCTCCCTGTTATTT 118
QY 120 ATGCAATCGAGTTCAATTTTGGGACTGCTTGGGAATGTCATCTGAGTGGTTCGGTACTCT 179
Db 119 ATGCAATTTCTTCCCTGGGATTTCCAGGCAATGCAGTAGTATACCACTTACATTT 178
QY 180 TCTGCATGAAGACTGGAACAGCAGCAATGTCATCTTTTAACCTTTTCCATCTCTGACT 239
Db 179 TCAAAATGAGACCTTGGAGAGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATC 238
QY 240 TTGCTTCCCTGTCACCTTCCCATCTGATGAAGATTTATGCCAAT---GATAAGGGGA 296
Db 239 TGCTGTATCTGACACGCTCCCTTCCGCTTACCTTCCATTTCAACCTGTATAGCA 358
QY 297 CCTATGAGATGTTCTCTGTATGAAGCAACCGATATGCTTTCACACCAACCTCTACACCA 356
Db 299 TCTTTGGAGATTTTCATGTTAAGTTTATCCGCTTACCTTCCATTTCAACCTGTATAGCA 358
QY 357 GCATCTCTCTCCACCTTTTCATAGCAGGAGATATCTGCTCATGAAGTACCTCTTCC 416
Db 359 GCATCTCTCTCCACCTTTTCATAGCAGGAGATATCTGCTCATGAAGTACCTCTTCC 416
QY 417 GAGAACACTTTCTACAAAAGAGGATTTGCCATTTTAACTCTGCTGGCTGGCT 476
Db 419 GCTGCTTTTCCATTCACAACTCGATGTCAGTTGTCAGCTTCCATTTCAACCTGTATAGCA 478
QY 477 TAGTGACCTTAGAAGTTCTACCAAGTCTGGAACCCCTGAAACCAATCTCATTTACAGCC 536
Db 479 TTTCACTGTAGCTGTCAATTCAGATGACCTTCTTGATCATCAACCAAGAGGACCAACA 538
QY 537 GCAGTAACCTGCACTGCAATGCAAGTCTGGAACCCCTGAAACCAATCTCATTTACAGCC 596
Db 539 GATCAGCTGTCTGACCTACAGTTCCG-----ATGAATCAATACTATTAAGTGGT 592
QY 597 TCTGCTGACTTTGTGGCTTCTTAATCTCTCTCTGATGTCCTCTCTACTACA 656
Db 593 ACAACCTGATTTGACTGCAACTACTTTCTGCTCCCTTGTGATGACACTTTGCT 652
QY 657 AGATGCTAGTCTTTAAAGAGGAGGAGCAGCAGCAAGCAACTGCCCTGGCAGTGGACA 716
Db 653 ATACACGATTTTACACACTCTGACCCATGGAAGTCAACCTGCAAGTGCCTTAAGCAGA 712
QY 717 AACCCCAACGCTGGTGGCTGGGTTGTGATCTCTCTACTCTTCACTCTCAACCCCTATC 776
Db 713 AAGCAGGAGGCTACCACTTCTGCTACTCTCTGATTTTACGATGTTTTTACCCCTTCC 772
QY 777 ATATCATGCGCAATTTGAGGATCGCCTCACGCTG 811
Db 773 ATATCTTGAGGGTCATTCGGATCGAATCTCGCCTG 807

```

RESULT 4  
US-10-270-144-3  
; Sequence 3, Application US/10270144  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000750CON

; CURRENT APPLICATION NUMBER: US/10/270.144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 9905  
; TYPE: DNA  
; ORGANISM: Human  
US-10-270-144-3

Query Match 8.2%; Score 126.6; DB 9; Length 9905;  
Best Local Similarity 49.9%; Pred. No. 1e-20;  
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

```

QY 60 CTTGTGAGATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCAATTT 119
Db 8367 CTTTGGAAATGGCAATGATGAATAATCCCACTCAAGATGCACCTACCTCCCTGTTATTT 8426
QY 120 ATGCAATCGAGTTCAATTTTGGGACTGCTTGGGAATGTCACCTCTGCTGTTTCGGCTACTCT 179
Db 8427 ATGCAATTTATCTTCTCTGCTGGGATTTCCAGGCAATGCAGTAGTATACCACTTACATTT 8486
QY 180 TCTGCATGAAGACTGGAACAGCAGCAATGTCATCTTTTAACTTTTCCATCTCTGACT 239
Db 8487 TCAAAATGAGACCTTGGAGAGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATC 8546
QY 240 TTGCTTCTCTGTCACCTTCCCATCTCTGATAAAGAGTTATGCCAAT---GATAAGGGGA 296
Db 8547 TGCTGTATCTGACACGCTCCCTTCCGCTGATTCATCTACTATGCGGAGGCAAACTGGA 8606
QY 297 CCTATGAGATGTTCTCTGTATGAAGCAACCGATATGCTTTCACACCAACCTCTACACCA 356
Db 8607 TCTTTGGAGATTTTCATGTTAAGTTTATCCGCTTTCAGCTTCCATTTCAACCTGTATAGCA 8666
QY 357 GCATCTCTCTCCACCTTTTCATAGCAGGAGATATCTGCTCATGAAGTACCTCTTCC 416
Db 8667 GCATCTCTCTCCACCTTTTCAGCAATCTCCGCTACTGCTGATCATTCACCAATGA 8726
QY 417 GAGACACTTTCTACAAAGAGGAAATTTGCCATTTTAACTCTGCTGGCTGCTGGGCT 476
Db 8727 GCTGCTTTTCCATTCACAAACTCGATGTCAGTTGTCAGCTGCTGCTGGTGGATCA 8786
QY 477 TAGTGACCTTAGAAGTTCTACCAAGTCTCACTTTTCATCAATCTGTCCTCCCAAGAGAGG 536
Db 8787 TTTCACTGCTAGTCTCATTCGAGTACCTTCTGATCATCAACCAAGAGGACCAACA 8846
QY 537 GCAGTAACCTGCACTGCAATGCAAGTCTTGGAAACCCCTGAAACCAATCTCATTTACAGCC 596
Db 8847 GATCAGCTGTCTGACCTCACCAGTTTCGG-----ATGAATCAATACTATTAAGTGGT 8900
QY 597 TCTGCTGACTTTTGGGCTTCTTAATCTCTCTCTGATGTCCTCTCTACTACA 656
Db 8901 ACAACCTGATTTTGGTGGGCTTCTTAATCTCTCTCTGATGTCCTCTCTACTACA 8960
QY 657 AGATGCTAGTCTTTAAAGAGGAGGAGCAGCAGCAAGCAACTGCCCTGGCAGTGGACA 716
Db 8961 ATACCAAGATTTACCACTCTGACCCATGGAAGTCAACCTGACAGCTGCTTTAAGCAGA 9020
QY 717 AACCCCAACGCTGGTGGCTGGGTTGTGATCTCTCTACTCTTCACTCTCAACCCCTATC 776
Db 9021 AAGCAGGAGGCTAACCACTTCTGCTACTCTCTGATTTTACGATGTTTTTACCCCTTCC 9080
QY 777 ATATCATGCGCAATTTGAGGATCGCCTCACGCTG 811
Db 9081 ATATCTTGAGGGTCATTCGGATCGAATCTCGCCTG 9115

```

RESULT 5  
US-10-023-775B-1  
; Sequence 1, Application US/10023775B  
; Publication No. US2003002282A1

```

; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (EP(CB) only)
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)
; APPLICANT: Fidock, Mark David
; TITLE OF INVENTION: No. US2003002282A1e1 Polypeptide
; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-023-775B-1

```

Query Match	8.1%	Score 125;	DB 9;	Length 1014;
Best Local Similarity	49.8%;	Pred. No. 7.3e-21;		
Matches 376;	Conservative	0; Mismatches 370;	Indels	9; Gaps
Qy	60	CTTGTGGAATTCGGTTGGCAACAGAGGCTATCTTGAATAAAGTACTACTCTCTCTCGATTTT	119	
Db	59	CTTTTGGAAATTCACATGATGAAACATCCCACTCAAGATGCACATACCTCGCTGTATT	118	
Qy	120	ATGCAATFCGAGTTTCATTTTGGACTCCTTTGGGAATGTCACCTGTGGTGTTCGGCTACCTCT	179	
Db	119	ATGCGATTATCTTCCCTCGTGGGATTTCCAGGCAATGCAGTAGTGATATCCACTTACATTT	178	
Qy	180	TCTCATGGAAGAAGCTGGAACAGCAGCAATGTCTATCTTTTAAACCTTTCCATCTCTGACT	239	
Db	179	TCAAAATGAGACCTTGAAGAGCAGCACCATCATATGCTGAACTGGCGCTGCACAGATC	238	
Qy	240	TTGCTTTTCCTGTGGACCTTCCCATCCTGNTAAGAGTTATGCCAAT---GATAAGGGGA	296	
Db	239	TGCTGTATCTTGACAGCTCCCTTCTGTAATCCTACTACTGCGAGTGGCGAAACTGGA	298	
Qy	297	CTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCA	356	
Db	299	TCTTTGGAGATTTCATGTGTAAGTTATCGGCTTCAGCTTCCATTTCACACCTGTATAGCA	358	
Qy	357	GCATCCTCTTCCCTCACCTTCATTTAGCATGACCGCATATCTGCTCATGANGTAGTACCCTTCC	416	
Db	359	GCATCCTCTTCCCTCACCTGTTTCAGCATCTTTCGCGTACTGTGTGATCATTCACCCAAATGA	418	
Qy	417	GAGAACACTTCTTACAAAAGAGGAATTTGCCATTTTAACTCGCTGGCTGTCTGGGCT	476	
Db	419	GCTGTTTTCATTTCACAAACTCGATGTGAGTTGTAGCCTGTGCTGTGGTGTGATCA	478	
Qy	477	TAGTGACCTTAGAAGTTCTTACCCATGCTCATCTTTTCATCAATTTCTGTGCCAAAGAGAGG	536	
Db	479	TTTCACTGGTAGCTGTGATTCGATGACCTTCTGTGATCATCAACACACAGGACCAACA	538	
Qy	537	CGAGTAAGTGCAGCTATGCAAGTTCTGTGAAACCTGAAACAAATCTCATTTTACAGCC	596	
Db	539	GATCAGCTGTCTGACCTCACCAAGTTCCG-----ATGAACTCAATACTATTAAGTGT	592	
Qy	597	TCTGCTGACTTTGTGTGGGCTTCTAATTCCTCTCTGTGATGTGGTCTTCTTACTACA	656	
Db	593	ACAACCTAATTTTGACTGCAACTACTTTCTGCTCCCTTGGTGATAGTGACACTTTTGTCT	652	
Qy	657	AGATGGTAGTCTCTTAAAGAGGAGGACCCAGCAGCAAGCAACTGCCCTGCCACTGGACA	716	
Db	653	ATACACAGATTATCCACACTCTGACCCATGGACTGCAAACTGACAGCTGCCTTAAGCAGA	712	
Qy	717	AAACCCCAACGCTGGTGTCTGGCGGTTGTGATCTTCTCTATACTCTTCAACCCCTATC	776	

[illegible]

Query Match	8.1%;	Score 125;	DB 9;	Length 1014;
Best Local Similarity	49.8%;	Pred. No. 7.3e-21;		
Matches	376;	Conservative	0;	Mismatches 370; Indels 9; Gaps
Qy	60	CTGTGTGAAATTCGGTTGGCAACAGAGGCTACTCTGAATAAGTACTACTCTCTCGCATTTT	119	
Db	59	CTTTTGGAAATTCGCATGTATGAAAGCATCCCACTCAAGATGCACACCTCCCTGTATT	118	
Qy	120	ATGCAATCGAGTTTCATTTTGGACTCGTTTGGAAATGTCACTGTGTGTTCGGCTACCTCT	179	
Db	119	ATGGCAATTCCTTCCTCGTGGGATTTCCAGGCAATACAGTAGTAGATATCCACTTACATTT	178	
Qy	180	TCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTGACT	239	
Db	179	TCAAAATAGCAACCTTGGAAAGCAGCAGCAATCATATGCTGAACCTGGCTGCACAGATC	238	
Qy	240	TTCCTTTCTGTGCACCTTCCCATCTCTGATAAAGAGTTTATGCCAAT---	296	GATAAGGGGA
Db	239	TGCTGTATCTGACACCGCTCCCTCTCTGATTCACCTACTATGCCAGTGCAGAAAAC	298	TGGA
Qy	297	CCTATGAGAGATGTTCTGTATATAGCAACCGATATGTGCTTACACACCACTCTACACCA	356	
Db	299	TCCTTTGGAGATTTTCATGTAAAGTTTATCGCTTTCAGCTTCCATTCACCTGTATAGCA	358	
Qy	357	GCATCCTCTTCCTCACATTTTCATTAGCATGAGACGGATATCTGTCATGAAGTACCCCTTCC	416	
Db	359	GCATCCTCTTCCTCACCTGTTTCAGCAATCTTCGCTACTGTGTGATCAITTCACCAATGA	418	
Qy	417	GAGAACACTTCTTACAAAAGAGGAATTTGCCATTTTAAATCTCGCTGGCTGTCTGGGCT	476	
Db	419	GCTGCTTTTCCATTCACAAAACTCGATGTGCAGTTGTAGCTGTGCTGTGGTGTGGATCA	478	
Qy	477	TAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTTCTGNTCCCAAGAAGAGG	536	
Db	479	TTTCACCTGGTAGCTGTCAATTCGGATGACCTTCTTGATGATCAATCAACCAACAGGACCA	538	
Qy	537	GCAGTAACCTGCATCGACTATGCAAGTTTCTGAAACCCCTGAACACAATCTCATTTACAGCC	596	
Db	539	GATCAGCCTGTCTCGACCTCACCAAGTTTCGG-----ATGAACCTCAATTAATAAGTGT	592	
Qy	597	TCTGCCTGACTTTGTGGGCTTCTTAATTCCTCTCTGTGATGTGCTTCTTCTACTACA	656	
Db	593	ACAACCTAATTTTGACTGCAACTACTTCTGCTCCCTTGGTGATAGTAGACACTTTGCT	652	
Qy	657	AGATGGTAGTCTCTTAAAGAGGAGGACCGACAGCAAGCAACTGCGCCTGCCACTGGACA	716	
Db	653	ATACCACAGATTATCCACACTCTGACCCATGGACTGCAAACTGACAGCTGCCTTAAAGCAG	712	
Qy	717	AAACCCCAACGCTGGTGTCTGGGGTGTGTGATCTTCTCTATCTCTTACAGACCCCTATC	776	

Query Match	7.7%	Score 119;	DB 10;	Length 1313;
Best Local Similarity	52.3%;	Pred. No. 2.5e-19;		
Matches 288;	Conservative 0;	Mismatches 260;	Indels 3;	Gaps 1;
QY	60	CTTGTGAGNATGGTGTGGCAACAGAGGCTATCTTCAATAAGTACTACCTCTCTGCAATTTT	119	
Db	605	CTTTTGGAAATTTGCACGTGATGAAAACATCCCACTCAAGATGCATACCTCCCTGTTATTT	664	
QY	120	ATGCAATCGAGTTCATTTTGGACTGCTTTGGAAATGTCACGTGTGTGGCTGGCTACCTCT	179	
Db	665	ATGGCAATATCTTCCCTGCGGATTTCCAGGCAATGCAGTAGTGATGCCACTTACATTT	724	
QY	180	TCTGCATGAAGAACTTGGAAACAGCAGCAATGCTATCTTTTAACTTTCCATCTCTGACT	239	
Db	725	TCAAATGAGACCTTGGAAAGAGCAGCACCACATATTATGCTGAACCTGGCCTGCACAGATC	784	
QY	240	TTGCTTTCCTGTGACCCCTTCCATCCTGTATAAGAGTTATGCCAAT---	296	
Db	785	TGCTGTATCTGACGACCCCTCCCTCTCTGATTCACCTACTATGCCAGTGGCGAAACTGGA	844	
QY	297	CCTATGGAGATGTTCTCTGTATGAAGCAACCGATATGTGCTTCACACCAACCTCTCACCA	356	
Db	845	TCTTTGGAGATTTCAATGTGTAGATTTATCCGCTTCAGCTTCATTTCAACCTGTATAGCA	904	
QY	357	GCATCTCTTCCCTCACTTTTCATAGCATGACGCGATATCTGCTCATGAAGTACCCCTTCC	416	
Db	905	GCATCTCTTCCCTCACTGTTTCAGCATCTTCGCTACTGTGTGATCATTCACCCCAATGA	964	
QY	417	GAGAACACTTCTTACAAAAGAGAAATTTGCCAATTTTAATCTCGCTGGCTGCTGGCCT	476	

Db 965 GCTGCTTTTCCATTACAAAACGATGTCAGTTGTAGCTGTGCTGTGGTGGATCA 1024  
QY 477 TAGTGACCTTAGAGTTCTACCCATGCTCACTTTTCATCAATTCGTCCCAAAAGAGG 536  
Db 1025 TTTCACTGGTAGCTGTCATTCGGATGACCTTCTTTGATCACATCAACCAAGGACCAACA 1084  
QY 537 GAGTAAGTGCATGCTATGATCAAGTGTCTGGAAACCTGAAACACAATCTCATTTACAGCC 596  
Db 1085 GATCAGCCTGTCTCGACCTCACCAGTTCGGATGAACTCAATACTATTAAAGTGTACAAAC 1144  
QY 597 TCTGCTGACT 607  
Db 1145 TAATTTGACT 1155

RESULT 7  
US-09-943-798-1  
: Sequence 1, Application US/09943798  
: Patent No. US20020065215A1  
: GENERAL INFORMATION:  
: APPLICANT: Glaxo Group Limited  
: TITLE OF INVENTION: Polypeptide  
: FILE REFERENCE: OG1021  
: CURRENT APPLICATION NUMBER: US/09/943,798  
: CURRENT FILING DATE: 2001-08-31  
: NUMBER OF SEQ ID NOS: 4  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 1  
: LENGTH: 831  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-943-798-1

Query Match 6.2%; Score 96; DB 10; Length 831;  
Best Local Similarity 49.4%; Pred. No. 8.8e-14;  
Matches 311; Conservative 0; Mismatches 310; Indels 9; Gaps 2;

QY 185 ATGAAGACTGGAGACAGCAGCAATGTCTATTTTAACTTTTCCATCTCTGACTTTGCT 244  
Db 1 ATGAGACCTTGGAGAGCAGCAGCATATGCTGAACCTGGCTGCACAGATCTGCTG 60  
QY 245 TTCTCTGTCACCTTCCCATCTCTGATTAAGAGTATGCTCAAT---GATAGGGGACCTAT 301  
Db 61 TATCTGACCACTCCCTCTCTGATGATCACTACTATGCGAGTGGCGAAACTGGATCTTT 120  
QY 302 GGAGATGTTCTCTGTATAAGCAACCGATATGTCTTACACCAACCTCTACACAGCATC 361  
Db 121 GGAGATTTTCATGTGAAGTTTATCCGCTTCAGCTTCCATTTCAACTGTATAGCAGCATC 180  
QY 362 CTCCTTCTCATTTCATAGATGGACCGATATCTGCTCATGAAGTACCTTTCCCGAGAA 421  
Db 181 CTCCTCTCACCTGTTTCAGCATCTCCGCTACTGTGTGATCATTCACCAATGAGCTGC 240  
QY 422 CACTTTCTACAAAAGAGGAATTTGCAATTTTAACTCGCTGGCTGTCTGGCCCTTAGTG 481  
Db 241 TTTTCATTACAAAACACTGATGCGAGTTGTAGCTGTGCTGTGGTGGATCATTTCA 300  
QY 482 ACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAATTTCTGCCAAAAGAGAGGCGAGT 541  
Db 301 CTGGTAGCTGTCTATTCGGATGACCTTCTTTGATCATCAACCAACAGGACCAACAGATCA 360  
QY 542 AACTGCATCGACTATGCAAGTTCTGGAACCCCTGAACACAATCTCATTTACAGCCTTCGC 601  
Db 361 GCCTGTCTCGACTCACCAGTTTCG-----ATGAACCTCAATACTATTAAAGTGTACAAAC 414  
QY 602 CTGACTTTTGTGGGCTTCTTAATCTCTCTCTGTGTGATGCTCTTCTTCTACTACAAGATG 661  
Db 415 CTGATTTTGACTGCACTACTTCTGCTCTCCCTTGGTGTAGTAGACACTTTGCTTATACC 474  
QY 662 GTAGCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCC 721  
Db 475 ACGATTATCCCACTCTGACCCACTGAGCTGCAAACTGACACCTGCCCTTAACACAGAAACA 534

QY 722 CAACGCTGTGCTCTCTGCGGTTGTGATCTCTCTATATCTCTACACCCCTATCATATC 781  
Db 535 CGAAGCTAACCAATCTGCTACTCTCTGCAATTTAGGATGTTTTTACCCCTCCATATC 594  
QY 782 ATGCGCAATTTGAGGATCGCTCACGCCG 811  
Db 595 TTGAGGCTATTCGGATCGAATCTCGCCTG 624

RESULT 8  
US-09-823-114-18  
: Sequence 18, Application US/09823114  
: Patent No. US20020061554A1  
: GENERAL INFORMATION:  
: APPLICANT: EVANS, CHRISTOPHER J.  
: KEITH, DUANE E.  
: TITLE OF INVENTION: OPIOID RECEPTOR GENES  
: NUMBER OF SEQUENCES: 25  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: MORRISON & FOERSTER  
: STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500  
: CITY: WASHINGTON  
: STATE: DC  
: COUNTRY: USA  
: ZIP: 20006-1888  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/823,114  
: FILING DATE: 29-Mar-2001  
: CLASSIFICATION: <Unknown>  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 09/148,351  
: FILING DATE: <Unknown>  
: ATTORNEY/AGENT INFORMATION:  
: NAME: MURASHIGE, KATE H.  
: REGISTRATION NUMBER: 29,959  
: REFERENCE/DOCKET NUMBER: 22000-20526.22  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202) 887-1500  
: TELEFAX: (202) 887-0763  
: TELEX: 90-4030 MRSNFOERSWSH  
: INFORMATION FOR SEQ ID NO: 18:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1805 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 10..1119  
: SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-823-114-18

Query Match 5.3%; Score 82.2; DB 10; Length 1805;  
Best Local Similarity 44.5%; Pred. No. 3.4e-10;  
Matches 379; Conservative 0; Mismatches 463; Indels 9; Gaps 1;

QY 85 GGCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCAATTTTGGACT 144  
Db 147 GCCCTCGGGCTCAAGTCAACCTCGTGGGGCTCTACCTGGCGCTGTGTGTCGAGGGCT 206  
QY 145 GCTTGGGAATGTCACCTGTGTGTGCGGTACCTCTTCTTGCATGAAGAACTGGAACAGCAG 204  
Db 207 CCTGGGAACCTGCTGTGTGTGTACGTACCTCTCAGGCACACCAAAATGAAGACGCCAC 266  
QY 205 CAATGTCTATCTTTTAACTTTCCATCTCTGACTTTGCTTTCTCTGTGACCCCTTCCCAT 264  
Db 267 CAATATTTTACATCTTTAACTGGCCCTGGCCGACACTCTGGTCTCTGCTGACGCTGCCCTT 326



265 CTTGATAAAGAGTTATGCCAATGATGAAGGACCTATGAGATGTTCTCTGTATAGCAA 324  
 327 CCAGGACAGGACATCCTCTGGGCTTCTGGCGTTTGGGAATGGCTTGCAGACAGT 386  
 325 CCGATATGCTTTCACCAACCACTCTACACGACATCTTCTTCCATCTTCAATAGCAT 384  
 387 CATTTGCCATGACTACTACACATGTTTACACGACCTTCCACCTAACTGCCATGAGTGT 446  
 385 GGACCGATATCTGCTCATGAAGTACCCCTTTCGAGAACACTTCTACAAAAGAGAAAT 444  
 447 GGATCGCTATGAGCATCTGCCACCCATCCGTCGCCCTCGACGTCGCCACCTCCAGCAA 506  
 445 TGCCATTTTAACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504  
 507 AGCCAGGCTGTCATGTTGGGCTTCTGGGCTTCTGGGCTTCTGGGCTTCTGGGCTTCT 566  
 505 CACTTTTCATCAATTTCTGCTCCAAAGAGAGGACGACGATGCTGCTGCTGCTGCTGCT 564  
 567 CATCATGGCTCGGACAGGCTCGAGGATGAAGAGATCGAGTGCCTGGTGGAGATCCCTAC 626  
 565 TGGAAACCTTGACACACATCTCATTTACAGCCTCTGCTGCTGCTGCTGCTGCTGCT 624  
 627 CCTCAGGATTAAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686  
 625 TCCCTCTCTGATGCTGCTTCTTCTACTACAAGATGCTGCTGCTGCTGCTGCTGCTGCT 684  
 687 CCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746  
 685 CCAGCAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744  
 747 CCTGCTCTGGGCTCCGAGAGAGGACCGGACCTGCGGCGGCTGCTGCTGCTGCTGCTGCT 806  
 745 TGTGATCTTCTTCTACTCTTCAACCTCATATATCATGCGCAATTTGAGGATCGCCTC 804  
 807 GGTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866  
 805 ACAGCTGGATAGTTGGCCACAGGATGTACAGAGAGGACATCAATCTATATACACACT 864  
 867 AGGGCTGGGCTTACAGCGGACGAGAGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
 865 GACAGGCTCTGGGCTTCTGACAGTGGCATCAATCCCATCTTCTTCTTCTTCTTCTTCT 924  
 924 -----GGCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977  
 925 AGACCATTACA 935  
 978 TGAGAACTTCA 988

RESULT 9  
 US-09-905-186A-9  
 ; Sequence 9, Application US/09905186A  
 ; Publication No. US2003008289A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kreek, Mary Jeanne  
 ; APPLICANT: LaForge, Karl Steven  
 ; TITLE OF INVENTION: Alleles of the Human Orphanin  
 ; TITLE OF INVENTION: FQ/No. US2003008289A1  
 ; TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon  
 ; FILE REFERENCE: 600-1-284N  
 ; CURRENT APPLICATION NUMBER: US/09/905,186A  
 ; PRIOR FILING DATE: 2001-10-19  
 ; PRIOR FILING DATE: US 60/218,205  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 1829  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-09-905-186A-9  
 Query Match 5.3%; Score 82.2; DB 9; Length 1829;

Best Local Similarity 44.5%; Pred. No. 3.4e-10;  
 Matches 379; Conservative 0; Mismatches 463; Indels 9; Gaps 1;  
 QY 85 GGCTATCTTGAATAGTACTACCTCTCTGCAATTTTATGCAATCGAGTTCAATTTTGGACT 144  
 Db 171 GCGGCTCGGGCTCAAGGTCAACCTCTGCGGCTCTACCTGCGCGTGTGTGCGAGGGCT 230  
 QY 145 GCTTGGGAATGCTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 204  
 Db 231 CCGTGGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290  
 QY 205 CAATGCTATCTTTTAACTTTTCACTCTGCACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 264  
 Db 291 CAATATTTTACATCTTTAACTTGGGCTGCGGACACACTGCTGCTGCTGCTGCTGCTGCTGCT 350  
 QY 265 CCGTGAATGAAGATTTATGCAATGATGAAGGACCTATGAGGATGCTTCTGTATTAAGCAA 324  
 Db 351 CCAGGACAGGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410  
 QY 325 CCGATATGCTTTCACACCAACCTCTACACGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 384  
 Db 411 CATTTGCCATTTGACTACTACTACATGTTTCCAGCAGACCTTCAACCTTAACCTGCGTGAAGTGT 470  
 QY 385 GGACCGATATCTGCTCAATGAAGTACCTTTCGAGAACACTTCTTACAAAAGAGAAAT 444  
 Db 471 GGATGCTATGAGCACTCTGCCACCCATCCGTCGCCCTCGACGCTCCGACGCTCCAGCAA 530  
 QY 445 TGCCATTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504  
 Db 531 AGCCGAGGCTGCTCAATGTTGGGCTCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590  
 QY 505 CACTTTTCATCAATTTCTGCTCCCAAAAGAGAGGCGGACGAGTAACTGATCGACTATGCAAGTTC 564  
 Db 591 CATCATGGGCTCGGACAGGCTCGAGGATGAAGAGATCGAGTGCCTGCTGCTGCTGCTGCTGCT 650  
 QY 565 TGGAAACCTTGACACAACTCTCAATTTTACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624  
 Db 651 CCTCAGGATTTACTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710  
 QY 625 TCTCTCTCTGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 684  
 Db 711 CCGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770  
 QY 685 CCAGCAGCAAGCAACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744  
 Db 771 CCGTCTCTGGGCTCCGAGAGAGGACCGGAACTGCGGCGCATCACTCGGCTGCTGCTGCTGCT 830  
 QY 745 TGTGATCTTCTCTATACTCTTCAACCCCTATCATATCATGCGCAATTTGAGGATCGCCTC 804  
 Db 831 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890  
 QY 805 AGCCCTGGATAGTTGGCCCAAGGATGTACAGAGAGGCGCATCAATCTATATACACACT 864  
 Db 891 AGGGCTGGGCTTCAAGCAGAGAGGAGAGTGGCGTGGCCATCTTCTGCGCTTCTGCGCTTCTGCG 947  
 QY 865 GACACGCGCTCTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 924  
 Db 948 -----GGCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001  
 QY 925 AGACCATTACA 935  
 Db 1002 TGAGAACTTCA 1012

RESULT 10  
 US-09-905-186A-10  
 ; Sequence 10, Application US/09905186A  
 ; Publication No. US2003008289A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kreek, Mary Jeanne  
 ; APPLICANT: LaForge, Karl Steven  
 ; TITLE OF INVENTION: Alleles of the Human Orphanin  
 ; TITLE OF INVENTION: FQ/No. US2003008289A1  
 ; TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon



us-09-891-138a-1.rnpb

675	CATCATGGGCTCGGCACAGGTTCGAGATTGAAGAGATCAGTGTCTGTGGAGATCCCTAC	734
QY	TGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCCTACTTGTTGGGCTTCTTAAT	624
Dd	735 CCTCAGGATTAATGGGCCCGGTCTTTGCCATTCGATCTTCCTCTCTCTTCATCGT	794
QY	TCCTCTCTCTGTGATGTGCTTTCTTACTACAGATGCTAGTCTTCTTTAAGAGGAGGAG	684
Dd	795 CCOCGTGCTCGTCAATCTGTCTGTCTACAGCCTCATGATTCGGGGCTCCGTGGAGTCCG	854
QY	685 CCAGCACAGAACAATCGCCCTGCCTCGACAAAAACCCCAAGCCTGTGGTCTCTGGGGGT	744
Dd	855 CTTGCTCTCGGGCTCCCAGAGAAAGAACCGGAACCTGGGGCCATCACTTCGGCTGGTGCT	914
QY	745 TGTGATCTTCTATACTCTTACACCCCTATCATATCATGCGCAATTTTAGGATCGCCTC	804
Dd	915 GGTGTAGTGGCTCTGTCTGTGGCTGTGTGGAGCCTGTCCAGGCTTCTCGTGTGGCCA	974
QY	805 ACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAATCTATATACACT	864
Dd	975 AGGCTGGGGTTTCAGCCGAGCAGCAGACTGCGBTGGCCATCTCGCTTCTTGCAC---	1031
QY	865 GACAGGCCCTCTGGCTTTCTGAACAGTGCATCAATCCCATCTTCTACTTCTCTCATGGG	924
Dd		
QY		1084

```

RESULT 13
US-09-905-186A-7
; Sequence 7, Application US/0905186A
; Publication No. US20030008289A1
; GENERAL INFORMATION:
; APPLICANT: Kreek, Mary Jeanne
; APPLICANT: LaForge, Karl Steven
; TITLE OF INVENTION: Alleles of the Human Orphanin
; TITLE OF INVENTION: FQ/No. US20030008289Allopeptin Receptor Gene, Diagnostic Method
; TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon
; FILE REFERENCE: 600-1-284N
; CURRENT APPLICATION NUMBER: US/09/905.186A
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/218,205
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-905-186A-7

```

[illegible]



ADDRESSER: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
 ADDRESSEE: STUART & OLSTEIN  
 STREET: 6 Becker Farm Road  
 CITY: Roseland  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07068  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/104,792  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/461,244  
 FILING DATE: 05-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ferraro, Gregory D.  
 REGISTRATION NUMBER: 36,134  
 REFERENCE/DOCKET NUMBER: 325800-445  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-994-1700  
 TELEFAX: 201-994-1744  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1586 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 431..1495  
 US-09-104-792-1

Query Match 5.2%; Score 80; DB 10; Length 1586;  
 Best Local Similarity 47.3%; Pred. No. 1.le-09;  
 Matches 276; Conservative 0; Mismatches 305; Indels 3; Gaps 1;

Qy	98	AAGTACTACCTCTCTGCAATTTTATGCAATCGAGTTCATTTTGGACTGCTGGGAATGTC	157
Db	533	AAGTTGCTCCTGCTGCTTTTATTGCTCTCTGTTTATTTCAGTCTCTCTGGGAAACAGC	592
Qy	158	ACTGTGGTTCGCGTACCTCTCTGATGAAGAACTGGAACAGCAGCAATGCTATCTT	217
Db	593	CTGGTCATCTCGGTCCTTGGGCTGCAAGAACTGAGGAGCATCAGATGTATACCTC	652
Qy	218	TTTAACCTTTCCATCTCTGACTTTGCTTCTGTGACCCCTTCCCATCTGTATAAAGAGT	277
Db	653	TTGAACCTGGCCCTGTGACCTGCTTTTGTCTCTCTCTCCCTTTCAGACCTA---C	709
Qy	278	TATGCCAATGATAGGGACCATATGGAGATGTTCTCTGTATAGCAACCGCATATGCTT	337
Db	710	TATCTGTGGACCATGCGGTGTTTGGGACTGTAATGTGCAAAAGTGGTGTGCTGGCTTTAT	769
Qy	338	CACACCAACCTCTACACGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	397
Db	770	TACATGGCTTCTACAGCAGCATGTTTTCATCACCCCTCATGAGTGTGGACAGGTACCTG	829
Qy	398	CTCATGAAGTACCCCTTTCCGAGAACACTTTCTACAAAGAAAGAAATTTGCCATTTTAAATC	457
Db	830	GCTGTGTCATGCGGTGTTTGGGACTGTAATGTGCAAAAGTGGTGTGCTGGCTTTAT	889
Qy	458	TCGCTGGCTGTCTGGCCCTTAGTGACCTTAGAAGTTCTACCATGCTCTCTCTCTCTCTCT	517
Db	890	TGCCTGGCAGTATGGCTAACCAGCATATGAGCTACCATCCCATTTGCTAGTCTTTACCAA	949
Qy	518	TCTGTCCCAAAAGAGAGGCGAGTAACTGTCATGCAAGTTCGTGGAAACCCCTGAA	577
Db	950	GTGGCCTCTGAAGATGGTGTCTACAGTGTATTATTTTACATTTTACAAATCAACAGACTTTGAAG	1009

Qy 578 CACAATCTCATTTACAGCCTCTGCTGACTTTGTTGGGCTTCTCTAATTTCTCTCTG 637  
 Db 1010 TGGAGATCTTCACCAACTTCAAAATGAACATTTTAGGCTTGTGATCCCATTCACCATC 1069  
 Qy 638 ATGTGCTTCTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAG 681  
 Db 1070 TTTATGTTCTGCTACATTTAAATCTCGACCATGCTGAAGAGGTG 1113

Search completed: April 2, 2003, 15:16:37  
 Job time : 170 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2003, 12:17:26 ; Search time 2695 Seconds  
(without alignments)  
9272.592 Million cell updates/sec

Title: US-09-891-138A-1

Perfect score: 1543

Sequence: 1 gctcctggcagagtttctg.....tgccataaataatcaatata 1543

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_htc: \*  
9: gb\_estl: \*  
10: gb\_est2: \*  
11: gb\_htc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rod: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	560	36.3	683	10	BB323771
2	495.8	32.1	520	9	AI663305
3	455	29.5	469	10	BB744515
4	438	28.4	458	10	BB746222
5	414	26.8	428	10	BB738743
6	403.8	26.2	422	10	BB847918

7	388.4	25.2	420	10	BB864882	
8	384.8	24.9	426	10	BB778587	
9	380.4	24.7	396	10	BB739482	
C	10	363.8	23.6	367	9	AI649254
11	357.6	23.2	636	10	BB645274	
12	354.2	23.0	416	10	BB846608	
13	309.8	20.1	327	10	BB845274	
14	296	19.2	877	12	BB220946	
15	294.6	19.1	333	10	BB220946	
16	283.2	18.4	323	10	BB220888	
17	279	18.1	316	10	BB225749	
18	275.2	17.8	326	10	BB500452	
19	261.4	16.9	285	10	BB327439	
C	20	255.8	16.6	633	17	AG083174
21	253.4	16.4	314	10	BB498575	
22	241	15.6	310	10	BB221521	
23	215.4	14.0	282	10	BB215653	
24	214	13.9	312	10	BB498898	
25	202	13.1	1101	17	CNS04VV3	
26	199.8	12.9	226	10	BB213317	
C	27	192.8	12.5	956	17	CNS028Y4
28	181	11.7	294	10	AW112068	
C	29	158.8	10.3	639	10	AW612141
30	151	9.8	1026	17	CNS051MY	
C	31	144.2	9.3	589	12	BF196066
32	141	9.1	1022	17	CNS04W90	
C	33	139.8	9.1	582	10	BE221739
34	137.4	8.9	139	9	AI021184	
35	90	5.8	638	9	AL675845	
36	89.4	5.8	641	14	BQ396255	
37	88.8	5.5	770	9	AJ452673	
38	84.6	5.5	2542	11	AK017378	
39	84.6	5.5	3001	11	AK005013	
40	80.4	5.2	546	12	BG552112	
41	78.4	5.1	952	14	BM917763	
42	78.4	5.1	1074	14	BM917063	
43	78.4	5.1	1167	14	BQ053936	
44	77.8	5.0	623	14	BQ038875	
45	76.8	5.0	672	10	BB638766	

#### ALIGNMENTS

RESULT 1 BB323771 683 bp mRNA linear EST 31-AUG-2001  
BB323771  
DEFINITION BB323771 RIKEN full-length clone B430012021 3', mRNA sequence.  
Mus musculus cDNA clone B430012021 3', mRNA sequence.

ACCESSION BB323771  
VERSION BB323771.2  
KEYWORDS GI:15411432  
SOURCE EST

ORGANISM house mouse.

REFERENCE

AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 683)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

JOURNAL Unpublished (2001)

COMMENT

On Jul 11, 2000 this sequence version replaced gi:9032085.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Thu Apr 3 10:42:32 2003

Fax: 81-45-503-9216  
 Email: genome-res@gs.riken.go.jp,  
 URL: <http://genome.gsc.riken.go.jp/>  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 The RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
 Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T.,  
 Ishii, Y. and Hayashizaki, Y.  
 Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct.  
 Func. Genomics* 2 pre, L72-L86 (2001)  
 please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
 further details.  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN,  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues

FEATURES  
source

1. `organism="Mus musculus"`  
`db_xref="taxon:10090"`  
`clone.lib="B430012021"`  
`male adipose"`  
`sex="male"`  
`tissue_type="adipose"`  
`dev_stage="4 days neonate"`  
`lab_host="DH10B"`  
`note="Site_1: SalI; Site_2: BamHI; cDNA library was`  
`prepared and sequenced in Mouse Genome Encyclopedia`  
`Project of Genome Exploration Research Group in Riken`  
`Genomic Sciences Center and Genome Science Laboratory`  
`in RIKEN. Division of Experimental Animal Research in Riken`  
`contributed to prepare mouse tissues. 1st strand cDNA`  
`primed with a primer (5'`  
`GAGACAGAGAGAGGATCCACAGCTCTTTTTTTTTTTVN 3'), cDNA`  
`prepared by using therolase thermo-activated reverse`  
`transcriptase and subsequently enriched for full-length`  
`casp-trapper. cDNA went through one round of normalization`  
`to Rot = 10.0 and subtraction to Rot = 229.0. Second`  
`strand cDNA was prepared with the primer adapter of`  
`sequence (5' GAGACAGAGATCTCCAGTAAATTAATATCCGCCGCC`  
`3'). cDNA was cleaved with XhoI and BamHI. Vector: a`  
`modified p Bluescript K3(+) after bulk excision from La`  
`FIG 1."`

BASE COUNT	226 a	125 c	117 g	215 t
ORIGIN				
Query Match		36.3%	Score 560;	DB 10; Length 683;
Best Local Similarity		98.7%	Pred. No. 8.2e-125;	
Matches 596; Conservative		0; Mismatches 5;	Indels 3; Gaps 3;	

943	GCTGATTAGTAAGTTCAGAC-AATACATTCAAG-TCCCTTACATCCCTTC-AGGACATGAGC	999
QY		
DDb		
66	GCTGATTAGTAAGTTCAGCCAAATACTTCAAGTTCCCTTTACATCTTCAAGGACATAAGT	125
QY		
1000	TGCTTGGATGCAAGTCTTCACTCAGCCAAAATGAGACACACTTGATAAACAGTGTCTGTGCAGT	1051
QY		

Db	126	TCCTGGATGCAGGTTTTCACCTACGCCAAATAGACACACTTGTATATAACAGTGTGTGCAGT	185
QY	1060	TTAGCTTTTAACTAAGTAAGCAACCATTTCTAGAGCTTTTAGCTTTTCCACCATCTCCCAACCC	1119
Db	186	TTAGCTTTTAAATTAAGTAAGCAACCATTTCTAGAGCTTTTAGCTTTTCCACCATCTCCCAACCC	245
QY	1120	CCAGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTCTCTGATTT	1179
Db	246	CCAGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTCTCTGATTT	305
QY	1180	TAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTA	1239
Db	306	TAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTA	365
QY	1240	AGAAGTGAACAAAAGAGAACAAATATCTCAATCTTTGGACACTTAGATCTCGAAATCTT	1299
Db	366	AGAAGTGAACAAAAGAGAACAAATATCTCAATCTTTGGACACTTAGATCTCGAAATCTT	425
QY	1300	GGAATTTTAAAGACCTCTTTTCTATCATAGTCTAAAGGAATACAAAGATAGTGTGCAA	1359
Db	426	GGAATTTTAAAGACCTCTTTTCTATCATAGTCTAAAGGAATACAAAGATAGTGTGCAA	485
QY	1360	ATGCTGAATGCATTTCAATTCAGTTCAGTTCGATGAAGCGTGTCTCGAAATAGTCTTATT	1419
Db	486	ATGCTGAATGCATTTCAATTCAGTTCAGTTCGATGAAGCGTGTCTCGAAATAGTCTTATT	545
QY	1420	TTTATTTCTGTAATATTAATAATTTATGTGAAAAATGAATATAATTCATGTACAACTTA	1479
Db	546	TTTATTTCTGTAATATTAATAATTTATGTGAAAAATGAATATAATTCATGTACAACTTA	605
QY	1480	GATTTCTATTTGAAAAATATATTTCTTGAAAAATAACTGCTGTGCTTAATAAATCAA	1539
Db	606	GATTTCTATTTGAAAAATATATTTCTTGAAAAATAACTGCTGTGCTTAATAAATCAA	665
QY	1540	TATA 1543	
Db	666	TATA 669	
RESULT 2			
LOCUS	AI663305	520 bp	linear
DEFINITION	uk27c10.y1 Sugano mouse kidney mklia Mus musculus cdna clone		EST 10-MAY-1999
KEYWORDS	IMAGE:1970226 5' similar to SW:P2YR_RAT P49651 P2Y PURINOCPTOR		
KEYWORDS	EST.		
KEYWORDS	AI663305		
KEYWORDS	AI663305.1	GI:4766888	
KEYWORDS	EST.		
KEYWORDS	house mouse.		
KEYWORDS	Mus musculus		
KEYWORDS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; 1 (bases 1 to 520)		
REFERENCE	Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Pen, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., R., E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.		
REFERENCE	The WashU-NCI Mouse EST Project 1999		
REFERENCE	Unpublished (1999)		
REFERENCE	Other ESTs: uk27c10.x1		
REFERENCE	Contact: Marra M/WashU-NCI Mouse EST Project 1999		
REFERENCE	Washington University School of Medicine		
REFERENCE	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		
REFERENCE	Tel: 314 286 1800		
REFERENCE	Fax: 314 286 1810		
REFERENCE	Email: mouseest@wustl.edu		
REFERENCE	This clone is available royalty-free through LNL; contact the		
REFERENCE	IMAGE Consortium (info@image.llnl.gov) for further information		
REFERENCE	MGI:986986		
REFERENCE	Seq primer: custom primer used		
REFERENCE	High quality sequence stop: 490.		
REFERENCE	Location/Qualifiers		
FEATURES			

## FEATURES







```

|||||
Db 369 TTTTATTCTGTAATATAAATTTATGTGAAATGAATATATTAATCAATGACACAT 428
Qy 1478 TAGATTTTCTATTTCAAAATTAATATTTCTT 1507
Db 429 TAGATTTTCTATTTGAAATTAATATTTCTT 458

|||||
RESULT 5
BB738743 428 bp mRNA linear EST 15-OCT-2001
LOCUS
DEFINITION
BB738743 RIKEN full-length enriched, 6 days neonate spleen Mus
BB738743
ACCESSION
BB738743.1 GI:16141748
VERSION
EST.
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 428)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Shibata,K., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Saito,R., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akaira,S., Tanaka,T., Tomaru,A., Toya,I.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
M., Konno,H., Okazaki,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
Carninci,P., Shibata,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. 428
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F430109C18"
/clone_lib="RIKEN full-length enriched, 6 days neonate
spleen"
/tissue_type="spleen"
/dev_stage="6 days neonate"
/note="vector: pSPORT1; Site.1: SalI; Site.2: NotI; This
clone is among a rearranged set of 15,247 clones from 11

```

```

embryo cDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5 embryos
, and E12.5 female mesonephros/gonad) and one newborn
ovary cDNA library. Average insert size 1.5 kb. All
source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
expression profiling of mid-gestational placenta and embryo
using a 15,000 mouse developmental cDNA microarray. 2000,
Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
Large-scale cDNA analysis reveals phased gene expression
patterns during preimplantation mouse development, 2000,
Development, 127: 1737-1749; (3) Genome-wide mapping of
unselected transcripts from extraembryonic tissue of
7.5-day mouse embryos reveals enrichment in the t-complex
and under-representation on the X chromosome, 1998, Hum
Mol Genet 7: 1967-1978."
BASE COUNT 153 a 59 c 72 g 144 t
ORIGIN
Query Match 26.8%; Score 414; DB 10; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.7e-89;
Matches 414; Conservative 0; Mismatches 0; Gaps 0;
Qy 1130 AGTACAAGCTGGTCCACATGATCAGAAAGCAGCTCTCTGTTCTGATTAGGTTATAC 1189
|||||
Db 1 AGTACAAGCTGGTCCACATGATCAGAAAGCAGCTCTCTGTTCTGATTAGGTTATAC 60
Qy 1190 CCAGAGTATGGAATAAAGCATGAGAAAGCATGACATCTTCACCTTAAGAACTGAAC 1249
|||||
Db 61 CCAGAGTATGGAATAAAGCATGAGAAAGCATGACATCTTCACCTTAAGAACTGAAC 120
Qy 1250 AAAGAGACAATAATTTCTCAATGTTTGACACCTTAGGATCTGAAATCTTGGAAATTTTA 1309
|||||
Db 121 AAAGAGACAATAATTTCTCAATGTTTGACACCTTAGGATCTGAAATCTTGGAAATTTTA 180
Qy 1310 AGACCTCTTTTCTATCATCTAGTCTAAAAGGAATACAAAGTAGCTGCAATGCTGAATG 1369
|||||
Db 181 AGACCTCTTTTCTATCATCTAGTCTAAAAGGAATACAAAGTAGCTGCAATGCTGAATG 240
Qy 1370 CATTTTCATCTGGTTCAGGTCGATAAGCGTCTTTCTGAAATAGTCTTATTTTATTTCTTG 1429
|||||
Db 241 CATTTTCATCTGGTTCAGGTCGATAAGCGTCTTTCTGAAATAGTCTTATTTTATTTCTTG 300
Qy 1430 TAATATTAAATTTATGTTGAAATGAATGAATAATCAATGATACAACTAGATTTTCTAT 1489
|||||
Db 301 TAATATTAAATTTATGTTGAAATGAATGAATAATCAATGATACAACTAGATTTTCTAT 360
Qy 1490 TTGAAATTAATTTCTTTGAAAAATAACTGCTGCTAAATAAATCAATATA 1543
|||||
Db 361 TTGAAATTAATTTCTTTGAAAAATAACTGCTGCTAAATAAATCAATATA 414

RESULT 6
BB847918 422 bp mRNA linear EST 26-NOV-2001
LOCUS
DEFINITION
BB847918 RIKEN full-length enriched, adult male kidney Mus musculus
BB847918 cDNA clone F530201f11 5', mRNA sequence.
ACCESSION
BB847918.1 GI:17086293
VERSION
EST.
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 422)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akaira,S., Tanaka,T., Tomaru,A., Toya,T.

```

TITLE  
JOURNAL  
COMMENT

Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.

FEATURES  
source

1. 422  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="F530201F1"  
/clone\_lib="RIKEN full-length enriched, adult male kidney"  
/sex="male"  
/tissue\_type="kidney"  
/dev\_stage="adult"  
/lab\_host="SOLR"  
/note="Site: 1: XhoI; Site: 2: SstI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGCGCGCGCACTCGAGTTTTTTTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5'  
GAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCCCCCCCCC 3'].  
cDNA was cleaved with XhoI and SstI."  
BASE COUNT 104 a 100 c 88 g 130 t  
ORIGIN

Query Match 26.2%; Score 403.8; DB 10; Length 422;  
Best Local Similarity 99.5%; Pred. No. 4.9e-87;  
Matches 405; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTCTGCGAGAGTTTCTGTCGAGACAGAGCGCGAGAGATGGCAGAGATTTATC 60  
|||||  
DB 16 GCTCTGCGAGAGTTTCTGTCGAGACAGAGCGCGAGAGATGGCAGAGATTTATC 75  
|||||  
QY 61 TTGTCGAGATTTGGTGGCAACAGAGGCTATCTTGAATTAAGTACTACCTCTCGATTTTA 120  
|||||  
DB 76 TTGTCGAGATTTGGTGGCAACAGAGGCTATCTTGAATTAAGTACTACCTCTCGATTTTA 135  
|||||  
QY 121 TCGAATCGAGTTCAATTTTGGACTCTTGGGAATGTCACCTGTGGTGTTCGGGTACCTCTT 180

Db 136 TCGAATCGAGTTCAATTTTGGACTCTTGGGAATGTCACCTGTGGTGTTCGGGTACCTCTT 195  
|||||  
QY 181 CTGCATGAAGAACGAGCAAGCAATGCTATCTTTTAAACCTTTCCATCTCTGACTT 240  
|||||  
DB 196 CTGCATGAAGAACGAGCAAGCAATGCTATCTTTTAAACCTTTCCATCTCTGACTT 255  
|||||  
QY 241 TGCCTTCTGTCGACCTTCCCATCTCCATCTGATAAGAGATTATGCCAATGATAGAGGACCTA 300  
|||||  
DB 256 TGCCTTCTGTCGACCTTCCCATCTCTGATAAGAGATTATGCCAATGATAGAGGACCTA 315  
|||||  
QY 301 TGGAGATGTTCTCTGATATAGCAACCGCATGTGCTTCACACCAACCTCTACACGAGCAT 360  
|||||  
DB 316 TGGAGATGTTCTCTGATATAGCAACCGCATGTGCTTCACACCAACCTCTACACGAGCAT 375  
|||||  
QY 361 CTTCTTCTCTCACTTTCTATTAGCATGGACCGCATATCTCTCATGAAGT 407  
|||||  
DB 376 CTTCTTCTCTCACTTTCTATTAGCATGGACCGCATATCTCTCATGAAGT 422  
|||||

RESULT 7  
BB864882 420 bp mRNA linear EST 27-NOV-2001  
LOCUS BB864882 RIKEN full-length enriched, RCB-1283 B16 melanoma cDNA Mus  
DEFINITION musculus cDNA clone G430047C11 5', mRNA sequence.  
ACCESSION BB864882 GI:17111092  
VERSION BB864882.1  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 420)  
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii  
Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okado, T.,  
Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,  
Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa  
A., Takahashi, F., Takaku-Akhiru, S., Tanaka, T., Tomaru, A., Toya, T.,  
Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.











Ishii, Y. and Hayashizaki, Y.  
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.  
Func. Genomics 2 pre, L72-L86 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES	source
e mouse tissues.	Location/Qualifiers
1. . 636	
/organism="Mus musculus"	
/db_xref="taxon:10090"	
/clone_lib="B430012021"	
/clone_lib="RIKEN full-length enriched, 4 days neonate male adipose"	
/sex="male"	
/tissue_type="adipose"	
/dev_stage="4 days neonate"	
/lab_host="DH10B"	
/note="Site.1: Salt; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAAGTCCAGAGGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGTTAATTAATCCCCCCCCCCCC 3'], cDNA was cleaved with XhoI and BamHI. Vector: a modified p Bluescript KS(+) after bulk excision from Lambda	

BASE COUNT	170 a	118 c	140 g	207 t	1 others
ORIGIN					
Query Match	23.2%	Score 357.6;	DB 10;	Length 636;	
Best Local Similarity	91.7%;	Pred. No. 7.7e-76;			
Matches 389;	Conservative	0;	Mismatches 34;	Indels 1;	Gaps 1;
Qy	1	GCTCCTGCACAGTTTCTGTGCGAGACAGAAGCCGACAGCAGAGATGGCACAGAATTATC	60		
Db	20	GCTCCTGCACAGTTTCTGTGCGAGACAGAAGCCGACAGCAGATGGCACAGAATTATC	79		
Qy	61	TGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA	120		
Db	80	TWTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGGATTTA	139		
Qy	121	TGCAATCAGATTCATTTTTGGACGCTTGGGAATGCTCACTGTGGTGTTCGGCTACCTCTT	180		
Db	140	TGCAATCAGATTCATTTTTGGACGCTTGGGAATGCTCACTGGGTGTTCGGGTACCTCTT	199		
Qy	181	CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAAACCTTTCCATCTCTGACTT	240		
Db	200	CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAAACCTTTCCATCTCTGACTT	259		
Qy	241	TGCTTTCTGTGCACCTTCCCATCCTGATAAAGAGTTATGCCAATGATAGGGGACCTA	300		
Db	260	TGCTTTCTGTGGCACCTTCCCATCCTGATAAAGAGTTTGGCAATGATAAGGGGACCTA	319		
Qy	301	TGGAGATGTTCTCTGTATAAGCAACCGGATATGCTTCCACCAACCTCTACACC - AGCA	359		
Db	320	TGGAGATGTTCTCTGTATAAGCAACCGGATATGCTTCCACCAACCTTTAATCCAAGCT	379		
Qy	360	TGCTTCTCTCTCACTTTTCATTAGCATGGCCGATATCTGCTCAAGTACCCCTTCGAG	419		
Db	380	TACTTTTACTACTTATTTAGCATGGACCGATATTTGTTTATGAAGTGCCCTTTTCG	439		
Qy	420	AAAA 423			
Db	440	AAAA 443			

RESULT 12	
BB846608	
LOCUS	
DEFINITION	

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

JOURNAL  
COMMENT

FEATURES source

BB846608 416 bp mRNA linear EST 26-NOV-2001  
BB846608 RIKEN full-length enriched, adult male kidney *Mus musculus*  
cDNA clone F5300031245', mRNA sequence.

BB846608  
BB846608.1 GI:17084983  
EST.

house mouse.  
Mus musculus  
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae; Mus.  
1 (bases 1 to 416)

AKIMURA, T., ARAKAWA, T., CARINCI, P., FUJINO, M., HANAGAKI, T., HAYASHI, N., HIRAMOTO, K., HIRAKAWA, T., HIRAZONE, T., IMOTANI, K., ISHII, Y., ITO, M., KAWA, J., KOJIMA, Y., KONNO, H., KOUDA, M., MATSUYAMA, T., NAKAMURA, M., NISHI, K., NOMURA, K., NUMASAKI, R., OKAZAKI, Y., OKIDO, T., OKAMURA, K., SAKAI, C., SAKAI, K., SAKAZUME, N., SASEKI, D., SATO, K., SHIBATA, R., SHINGEGAWA, A., SHIRAKI, T., SOGABE, Y., SUZUKI, H., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKIHARA, S., TANAKA, T., TOMARU, A., TOYA, T., WAKABAYASHI, A., YASUNISHI, A., YAMAMOTO, M. and HAYASHIZAKI, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-2 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>,  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
, M., Konno, H., Okazaki, Y., Muramatsu, Y., Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10): 1617-1630 (2000)  
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Okawa, K., Tanaka, T., Watsuura  
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system - 394 - formal sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.  
Location/Qualifiers  
1. .416

```
/organism="Mus musculus"  
/db_xref="taxon:10090"  
/clone="F53003I24"  
/clone_lib=RIKEN full-length enriched, adult male kidney  
/sex="male"  
/tissue_type="kidney"  
/dev_stage="adult"  
/lab_host="SOLR"  
/note="Site_1: XhoI; Site_2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',  
GAGAGAGAGCGCCGCAACTCGAGTGTCTTTTTTTTTTTVN 3'], cDNA was prepared by using trihalose thermo-activated reverse
```



```

QY 1523 GTGCCCTAAATAAATCAATATA 1543
      |||||||
Db 301 GTGCCCTAAATAAATCAATATA 321

RESULT 14
BG402029 877 bp mRNA linear EST 12-MAR-2001
LOCUS 602466748F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594810 5',
DEFINITION mRNA sequence.
ACCESSION BG402029 GI:13295477
VERSION BG402029.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1336 row: p column: 11
High quality sequence stop: 542.

FEATURES
source
Location/Qualifiers
1..877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4594810"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccattagggc); Site 2: SfiI (ggccattagggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 198 a 221 c 199 g 259 t
ORIGIN
Query Match 19.2%; Score 296; DB 12; Length 877;
Best Local Similarity 74.6%; Pred. No. 6.4e-61;
Matches 412; Conservative 0; Mismatches 135; Indels 5; Gaps 3;

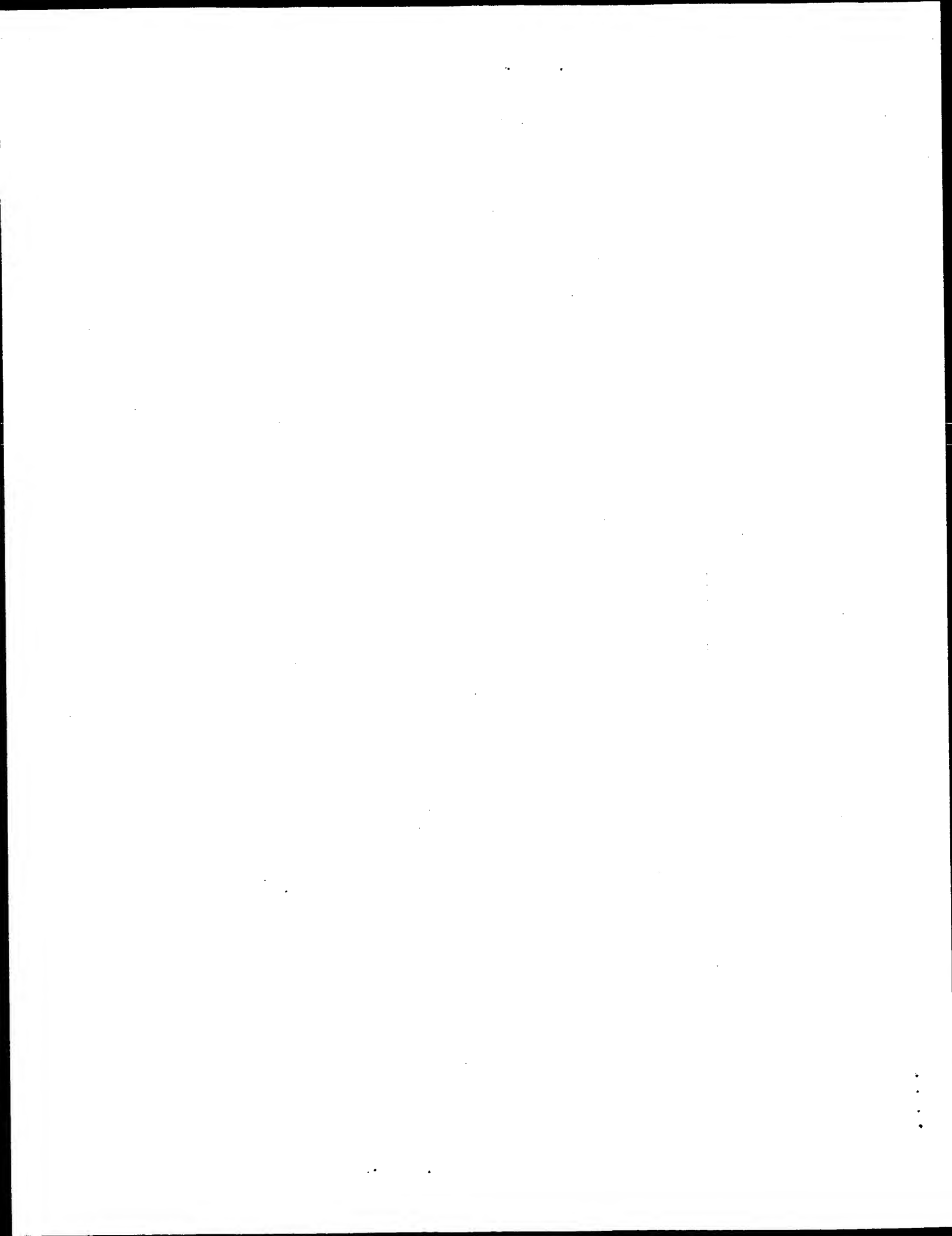
QY 39 GCAGATGGCAGCAATTTATCTTGTGAGAACTGGTGGCAACAGAGGCTATCTTGAATA 98
      |||||||
Db 108 GGATATGGCATGGAATGCAACTTGCAGAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 167
      |||||||
QY 99 AGTACTACCTCTCTGCATTTTATTCATTCGATTCATTTTGGAGCTCTTGGGAATGCA 158
      |||||||
Db 168 AGTACTACCTTTCCATTTTATGGGATGTAGTTCGTGGGAGTCTTGGAAATACCA 227
      |||||||
QY 159 CTGGGTTGGCTGCTCTCTGATGAGAACTGGCAACAGCAGCAATCTCTATCTTT 218
      |||||||
Db 228 TTGTGTTTACGGCTACATCTCTCTCTGAAGAACTGGCAACAGCAGTAAATTTATCTCT 287
      |||||||
QY 219 TTAACCTTTTCCATCTCTGACTTTTCTCTCTGTCGACCTTCCCATCTCTGATAAAGAGTT 278
      |||||||
Db 288 TTAACCTCTCTCTCTGACTTGTCTCTCTGTCGACCTTCCCATCTCTGATAAAGAGTT 347
      |||||||

```

279 ATGCCAATGATAGGAGGACCTATGGAGATGTTCTCTGTATAACCAACCGATATGCTCTTC 338  
|||||  
348 ATGCCAATGGAACCTGGATATATGGAGACGCTGCTGCATAAGCAACCGATATGCTCTTC 407  
|||||  
339 ACACCAACCTCTACACAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398  
|||||  
408 ATGCCAACCTCTATACACGCAATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 467  
|||||  
399 TCATGAAGTACCTCTCTCCGAGAACACCTTTCTACAAAAGAGGAATTTGCCATTTTAATCT 458  
|||||  
468 TAAATAGTATCTCTCTCCGAGAACACCTTTCTGCAAGAAAG--AGTTGCTATTTTAATCT 525  
|||||  
459 CGCTGGCTGTCTGGGCTCTTAGTGACCTTAGTAGAGTTTACCCATGCTCTCTCTCTCTCT 518  
|||||  
526 CCTGCC--ATTTGGGTTTACTTAACCTTAGAGTTTACCTCTCTCTCTCTCTCTCTCTCT 583  
|||||  
519 CTGTGCAAAAGAGAGGCGAGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578  
|||||  
584 CTGTATTATTAACCTGCAATGGCG--CCCTGTAAGGATTTTGAAGGCTGCTGAGACCCCACT 642  
|||||  
579 ACAATCTCATTT 590  
|||||  
643 CCCACCTCTTTT 654

RESULT 15  
BB254869 333 bp mRNA linear EST 06-JUL-2000  
LOCUS BB254869 RIKEN full-length enriched, 7 days neonate cerebellum Mus  
DEFINITION musculus cDNA clone A730064B21 3', mRNA sequence.  
ACCESSION BB254869  
VERSION BB254869.1 GI:8947615  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 333)  
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,  
P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,  
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,  
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,  
Y., Shigenoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,  
T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yananaka,I.,  
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,  
M., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Konno,H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
N. Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Thermotabilization and thermoactivation of full length  
trehalase and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95(2): 520-524 (1998)  
Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,  
Y. and Hayashizaki,Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5): 463-470 (1999)  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 13:08:42 ; Search time 39 Seconds  
(without alignments)  
1083.088 Million cell updates/sec

Title: US-09-891-138a-2

Perfect score: 1650

Sequence: 1 MAQNSCENWLAETAILNKY.....REMLTSKPROVFKSLTSFRT 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq.101002.\*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1650	100.0	317	23	AAU74904
2	1238.5	75.1	334	21	AAU71308
3	1238.5	75.1	334	21	AAH02842
4	1238.5	75.1	334	23	ABB90381
5	1238.5	75.1	379	23	AAE15633
6	1226.5	74.3	334	18	AAW19854
7	1219.5	73.9	334	18	AAW22732
8	1216.5	73.7	387	22	AAU31029
9	1014.5	61.5	258	21	AAH45376
10	474	28.7	373	22	AAE04391

11	474	28.7	373	23	AAU10983	Purinergic recepto
12	474	28.7	373	23	AAU10984	Purinergic recepto
13	467	28.3	373	23	AAU10985	Purinergic recepto
14	426	25.8	337	22	AAU04375	Human G-protein co
15	426	25.8	337	23	AAU15399	Human G-protein co
16	426	25.8	337	23	ABH81902	Human G-protein co
17	426	25.8	337	23	ABH83819	Human G-protein co
18	426	25.8	337	23	ABH83819	Human G-protein co
19	426	25.8	337	23	ABH83819	Human G-protein co
20	426	25.8	337	23	ABH83819	Human G-protein co
21	426	25.8	337	23	ABH83819	Human G-protein co
22	426	25.8	337	23	ABH83819	Human G-protein co
23	426	25.8	337	23	ABH83819	Human G-protein co
24	426	25.8	337	23	ABH83819	Human G-protein co
25	409	24.8	259	21	AAH45375	Human G-protein co
26	386.5	23.4	537	23	AAU74538	Gene 37 human secr
27	379.5	23.0	336	22	AAH80971	Human P2Y purinoce
28	373.5	22.6	373	23	AAE20604	Human NGPC854 #2
29	366	22.2	276	23	ABH83818	Mus musculus GPCR
30	363.5	22.0	377	22	AAE04392	Human P2Y-like rec
31	363.5	22.0	377	22	AAE01143	Human P2-purinergi
32	363.5	22.0	377	22	AAE01144	Human purinergic r
33	353	21.4	355	19	AAW48722	Seven transmembran
34	353	21.4	355	21	AAH21692	Human V28 seven tr
35	353	21.4	355	21	AAH21692	Human 7TM receptor
36	353	21.4	355	21	AAH21693	Human G-protein-co
37	353	21.4	355	21	AAH21693	Human mutant G-pro
38	353	21.4	355	21	AAH21693	Human CX3C chemoki
39	353	21.4	355	21	AAH21693	Human 7 transmembr
40	353	21.4	355	21	AAH21693	Human 7 transmembr
41	353	21.4	355	21	AAH21693	Protein CX3CRL dif
42	353	21.4	355	21	AAH21693	Human P2Y4 recepto
43	353	21.4	355	21	AAH21693	Human P2Y4 pyrimid
44	352	21.3	365	18	AAW23606	
45	352	21.3	365	22	AAE04391	

## ALIGNMENTS

RESULT 1  
AAU74904  
ID AAU74904 standard; Protein: 317 AA.  
AC AC  
AAU74904;  
09-APR-2002 (first entry)  
DE Amino acid sequence of mouse G-protein coupled receptor TGR18 protein.  
KW Mouse; G-protein coupled; receptor; GPCR; TGR18; kidney disease;  
KW signal transduction modulator; cerebral cavernous malformation;  
KW hyperlipidemia; obesity; dyslexia; cardiac myxoma; renal failure;  
KW nephritis; hypertension; liver disease; cirrhosis; blood disorder;  
KW spleen-associated disorder; immune disorder.  
XX Mus sp.  
XX WO200200719-A2.  
XX PN  
XX PD 03-JAN-2002.  
XX PF 25-JUN-2001; 2001WO-US20363.  
XX PR 23-JUN-2000; 2000US-213461P.  
XX PA (TULA-) TULARIK INC.  
XX PI Lin DC, Zhao J, Chen J, Cutler G;  
XX DR WPI; 2002-147880/19.  
XX N-PSDB; ABK12957.

PT	New G-protein coupled receptor polypeptides, useful for identifying	OS	Homo sapiens.
PT	modulators of signal transduction for treating kidney disease,	XX	
PT	hyperlipidemia, obesity, dyslexia and cardiac myxoma	PN	WO200031258-A2.
XX		XX	
PS	Claim 33; Page 59; 78pp; English.	PD	02-JUN-2000.
XX		XX	
CC	The present invention relates to a new G-protein coupled receptor (GPCR)	XX	13-OCT-1999; 99WO-US23687.
CC	polypeptide comprising greater than 70% amino acid sequence identity to	XX	98US-0109213.
CC	the amino acid sequence of human GPCRs TGR62, TGR21, TGR130.1, TGR130.2,	PR	99US-0120416.
CC	human TGR213 or TGR92, 80% amino acid sequence identity to mouse TGR18	PR	99US-0121852.
CC	or 90% amino acid sequence identity to human novel edg receptor protein,	PR	99US-0123946.
CC	as defined in the specification. The GPCR covalently linked to a solid	PR	99US-0123949.
CC	phase is useful for identifying a compound that modulates signal	PR	99US-0136436.
CC	transduction. The identified compounds are useful for treating	PR	99US-0136437.
CC	kidney disease, cerebral cavernous malformations, hyperlipidemia,	PR	99US-0136439.
CC	obesity, dyslexia and cardiac myxoma. The molecules of the invention are	PR	99US-0136567.
CC	useful for diagnosing disorders or conditions such as kidney-related	PR	99US-0137127.
CC	conditions or diseases such as renal failure, nephritis, nephrotic	PR	99US-0137131.
CC	syndrome, asymptomatic urinary abnormalities, renal tubule defects,	PR	99US-0141448.
CC	hypertension and nephrolithiasis, liver-related disease or condition	PR	99US-0156555.
CC	e.g. cirrhosis, infiltrations, lesions, functional disorders and jaundice	PR	99US-0156633.
CC	and spleen-associated disorders or conditions e.g. splenic enlargement,	PR	99US-0156634.
CC	immune disorders, blood disorders and others. Modulation of the	PR	99US-0156653.
CC	polypeptide of the invention is useful to treat or prevent any of the	PR	99US-0157280.
CC	above conditions or diseases. The present amino acid sequence represents	PR	99US-0157281.
CC	the mouse GPCR TGR18 protein of the invention. This sequence is one of	PR	99US-0157282.
CC	seven novel G protein coupled receptors of the invention (AAU74904-	PR	99US-0157293.
CC	AAU74911).	PR	99US-0157294.
XX		PR	99US-0416760.
SQ	Sequence 317 AA;	PR	99US-0417044.
Query Match 100.0%; Score 1650; DB 23; Length 317;			
Best Local Similarity 100.0%; Pred. No. 1.3e-157;			
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MAQNLSCEWLAETAILNKYLSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNYLNF 60	XX	(AREN-) ARENA PHARM INC.
Db	1 MAQNLSCEWLAETAILNKYLSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNYLNF 60	XX	Chen R, Dang HT, Liaw CW, Lin I;
QY	61 LSISDFALCTPLILIKSYANDKTYGDLVLCISNRYVLTNLTSLTFLTFISMDRYLLM 120	XX	WPI: 2000-400068/34.
Db	61 LSISDFALCTPLILIKSYANDKTYGDLVLCISNRYVLTNLTSLTFLTFISMDRYLLM 120	XX	N-PSDB; AAD01135.
QY	121 KYPREHFLOKKEFALLISLAWALVTLEVLPLMTFTNSVPKEGSCIDYASSGNPEHN 180	PT	Novel human orphan G protein-coupled receptors and the encoding cDNAs
Db	121 KYPREHFLOKKEFALLISLAWALVTLEVLPLMTFTNSVPKEGSCIDYASSGNPEHN 180	PT	for use in the identification of G protein-coupled receptor agonists -
QY	181 LYSICLTLGLFLPLSVCMCFYFKMVVFLKRRSQQQATALPLDKPQRLVVLAVVIFSI 240	XX	Claim 70; Page 87-88; 102pp; English.
Db	181 LYSICLTLGLFLPLSVCMCFYFKMVVFLKRRSQQQATALPLDKPQRLVVLAVVIFSI 240	XX	The present amino acid sequence is the hCHN10, an endogenous human
QY	241 FPPYHIMNRLRASLDSPGCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYREM 300	CC	orphan G protein-coupled receptor (GPCR), expressed in kidney and
Db	241 FPPYHIMNRLRASLDSPGCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYREM 300	CC	thyroid. The hCHN10 cDNA was identified using the human EST (expressed
QY	301 LISKFRQYFKSLTSFT 317	CC	sequence tag) 1365839 as a probe. The orphan GPCR of the invention, like
Db	301 LISKFRQYFKSLTSFT 317	CC	all GPCRs has seven transmembrane alpha helices with an extracellular
RESULT 2			
AAAY71308	Query Match 75.1%; Score 1238.5; DB 21; Length 334;		
ID	AAAY71308 standard; Protein: 334 AA.	Best Local Similarity 72.2%; Pred. No. 3.6e-116;	
XX		Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;	
AC	AAAY71308;		
XX	02-NOV-2000 (first entry)		
DT			
XX	Human orphan G protein-coupled receptor hCHN10.		
DE			
XX	Human; orphan G protein-coupled receptor; GPCR; hCHN10; drug screening;		
KW	transmembrane receptor; expressed sequence tag; EST; signal cascade.		
XX			
QY	1 MAQNLSCEWLAETAILNKYLSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNYLNF 60	QY	1 MAQNLSCEWLAETAILNKYLSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNYLNF 60
Db	5 MAWNATCKNWLAAEALEKYLIFIFGIEFVGVGLNTIVVGYIFSLKNNSSNYLNF 64	Db	5 MAWNATCKNWLAAEALEKYLIFIFGIEFVGVGLNTIVVGYIFSLKNNSSNYLNF 64
QY	61 LSISDFALCTPLILIKSYANDKTYGDLVLCISNRYVLTNLTSLTFLTFISMDRYLLM 120	QY	61 LSISDFALCTPLILIKSYANDKTYGDLVLCISNRYVLTNLTSLTFLTFISMDRYLLM 120
Db	65 LSISDFALCTPLILIKSYANDKTYGDLVLCISNRYVLTNLTSLTFLTFISMDRYLLM 124	Db	65 LSISDFALCTPLILIKSYANDKTYGDLVLCISNRYVLTNLTSLTFLTFISMDRYLLM 124



QY 121 KYPREHFLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 180  
 DB 125 KYPREHLLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 184  
 QY 181 LIYSCLTLGLFLPLSVKMFYKMYVFLKRRSQQAATLPLDKPQRLVLAIVFISIL 240  
 DB 185 LIYSCLTLGLFLPLSVKMFYKMYVFLKRRSQQAATLPLDKPQRLVLAIVFISIL 244  
 QY 241 FTPYHMRNLRIASRLDSWPO-GCTQKAISIVTLTRPLAFLNSAINPIFYFLMGDHYRE 299  
 DB 245 FTPYHMRNRVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYLLGDHFRD 304  
 QY 300 MLISKFRQYFKSLTSF 315  
 DB 305 MLMNQLRHNFKSLTSF 320

RESULT 3  
 AAB02842  
 ID AAB02842 standard; Protein; 334 AA.  
 AC AAB02842;  
 XX  
 DT 22-AUG-2000 (first entry)  
 XX  
 DE Human G protein coupled receptor hCHN10 protein SEQ ID NO:38.  
 XX  
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;  
 KW identification; agonist; screening; therapeutic; pharmaceutical;  
 KW mutant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200022131-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 13-OCT-1999; 99WO-US24065.  
 XX  
 PR 13-OCT-1998; 98US-0170496.  
 PR 12-NOV-1998; 98US-0108029.  
 PR 20-NOV-1998; 98US-0109213.  
 PR 27-NOV-1998; 98US-0110060.  
 PR 16-FEB-1999; 99US-0120416.  
 PR 26-FEB-1999; 99US-0121852.  
 PR 12-MAR-1999; 99US-0123944.  
 PR 12-MAR-1999; 99US-0123945.  
 PR 12-MAR-1999; 99US-0123946.  
 PR 12-MAR-1999; 99US-0123948.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 12-MAR-1999; 99US-0123951.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 30-JUN-1999; 99US-0137567.  
 PR 27-AUG-1999; 99US-0141448.  
 PR 03-SEP-1999; 99US-0151114.  
 PR 29-SEP-1999; 99US-0152524.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156555.  
 PR 29-SEP-1999; 99US-0156634.  
 XX  
 PA (AREN-) ARENA PHARM INC.  
 XX  
 PI Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;  
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;  
 XX  
 DR WPI; 2000-317986/27.  
 DR N-PSDB; AAA46036.  
 XX  
 PT Non-endogenous, human G protein-coupled receptors for screening

PT receptor, inverse or partial agonists useful as therapeutic agents -  
 XX Example 1; Page 117-118; 187pp; English.  
 XX The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCR), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,  
 CC inverse agonists or partial agonists for use as pharmaceutical agents,  
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX  
 XX Sequence 334 AA;

Query Match 75.1%; Score 1238.5; DB 21; Length 334;  
 Best Local Similarity 72.2%; Pred. No. 3.6e-116;  
 Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;  
 QY 1 MAONLSCENWLAETAILNKYYLSAFYAIEIFLGLGNVTVVFGYLFCKMKNSSNYLFN 60  
 DB 5 MAWNAATCKNMLAAEALEKYYLSIFGYIEFVGVGLNTIVVGYIFSLKNNSSNYLFN 64  
 QY 61 LISIDFAFLCTLPILIKSYANDKGTGYDVLICISNRYLHTNLYTSLFTFISMDRYLLM 120  
 DB 65 LSVSDLAFLCTLPMLIRSYANGNWIYDVLICISNRYLHANLYTSLFTFISIDRYLII 124  
 QY 121 KYPREHFLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 180  
 DB 125 KYPREHLLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 184  
 QY 181 LIYSCLTLGLFLPLSVKMFYKMYVFLKRRSQQAATLPLDKPQRLVLAIVFISIL 240  
 DB 185 LIYSCLTLGLFLPLSVKMFYKMYVFLKRRSQQAATLPLDKPQRLVLAIVFISIL 244  
 QY 241 FTPYHMRNLRIASRLDSWPO-GCTQKAISIVTLTRPLAFLNSAINPIFYFLMGDHYRE 299  
 DB 245 FTPYHMRNRVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYLLGDHFRD 304  
 QY 300 MLISKFRQYFKSLTSF 315  
 DB 305 MLMNQLRHNFKSLTSF 320

RESULT 4  
 ABB90381  
 ID ABB90381 standard; Protein; 334 AA.  
 AC ABB90381;  
 XX  
 DT 24-MAY-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 2757.  
 XX  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotrophic; antidiabetic; antiinflammatory; antitumor;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200190304-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-US16450.  
 XX  
 PR 19-MAY-2000; 2000US-205515P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX

XX  
KW  
Human: G-protein coupled receptor-3; GCRC-3; therapy; cancer; cancer; stroke;

1 MAQNLSCENWLATEAILNKYLSAFYAIEFIFGLLGNVTYVFGYLFCKMKNWNSSNVYLEN 60

1 MAQNLSCENWLATEAILNKYLSAFYAIEFIFGLLGNVTWVFGYLFCKMKNWSSNVYLEN 60  
QY  
109

```

QY 61 LSISDAFLCTPLILIKSYANDKGTGYDVLICISNRVYLVHTNLXTSLTLFTFISMDRYLLM 120
Db 110 LSVSDLAFLCTPLMIRSYANGNWIYGDVLCISNRVYLVHANLYTSILFTFISIDRYLLI 169
QY 121 KYPREHFLQKKEFAILISIAVWALVTLVPLMLTFINSVPKEEGSNCDIYASSGNPEHN 180
Db 170 KYPREHLLQKKEFAILISIAVWVLTLELLPLINPVITDNGTTCNDFASSGDPNPN 229
QY 181 LIYSICLTLGLFLPLSVCMCFYKVMVFLKRSQOQATALPLDKPQRLVLAUVIFSL 240
Db 230 LIYSWCLTLGLFLPLSVCMCFYKVMVFLKRSQOQATALPLDKPQRLVLAUVIFSVL 289
QY 241 FTPYHMRNRIASRLDSWPQ-GCTQKAIKSIYTLTRPLAFNLNSAINPIFYFLMGDHYRE 299
Db 290 FTPYHVMNRVRIASRLGSKWQYQCTQVINSFVITRPLAFNLNSVINPVFLLGDHFRD 349
QY 300 MLISKFRQYKSLTSF 315
Db 350 MLMNQLRHNFKSLTSF 365

RESULT 6
AAW19854
ID AAW19854 standard; Protein; 334 AA.
XX AC AAW19854;
XX DT 11-SEP-1997 (first entry)
XX DE Human purinergic receptor P2U2.
XX KW P2U2 receptor; purinergic receptor; diagnosis; therapy.
XX OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 1..23
FT Domain /label= N-terminal_domain
FT Domain 24..49
FT Domain /label= TMI
FT Domain /note= "transmembrane domain I"
FT Domain 50..60
FT Domain /label= ICDI
FT Domain /note= "intracellular domain I"
FT Domain 61..82
FT Domain /label= TMII
FT Domain /note= "transmembrane domain II"
FT Domain 83..99
FT Domain /label= ECDI
FT Domain /note= "extracellular domain I"
FT Domain 100..119
FT Domain /label= TMIII
FT Domain /note= "transmembrane domain III"
FT Domain 120..141
FT Domain /label= ICDII
FT Domain /note= "intracellular domain II"
FT Domain 142..161
FT Domain /label= TMDIV
FT Domain /note= "transmembrane domain IV"
FT Domain 162..183
FT Domain /label= ECDII
FT Domain /note= "extracellular domain II"
FT Domain 184..207
FT Domain /label= TMDV
FT Domain /note= "transmembrane domain V"
FT Domain 208..233
FT Domain /label= ICDIII
FT Domain /note= "intracellular domain III"
FT Domain 234..256
FT Domain /label= TMDVI
FT Domain /note= "transmembrane domain VI"
FT Domain 257..276

```

```

FT /label= ECDIII
FT /note= "extracellular domain III"
FT 277..300
FT /label= TMDVII
FT /note= "transmembrane domain VII"
FT 301..334
FT /label= C-terminal_domain
XX
PN WO9720045-A2.
XX
XX 05-JUN-1997.
XX
XX 08-NOV-1996; 96WO-US18175.
XX
XX 15-NOV-1995; 95US-0559524.
XX 15-NOV-1995; 95US-0006782.
XX
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Conley PB, Jantzen H;
XX
XX WPI; 1997-310601/28.
XX N-PSDB; AAT71900.
XX
XX New isolated purinergic receptor sub-type - used to develop
XX products for diagnosis and therapy, e.g. for screening for agonists
XX and antagonists which can modulate activation
XX
XX Claim 1; Fig 1A-B; 36pp; English.
XX
XX P2U2 receptor (AAW19854) is a novel human purinergic receptor
XX subtype that is abundantly expressed in kidney and in many cell
XX lines of megakaryocytic or erythroleukemic origin and which is
XX activated by ATP, UDP, UTP and UDP. Its amino acid sequence was
XX deduced from a cDNA clone derived from DAMI (ATCC CRL 9792) cells.
XX P2U2 and its polypeptides can be expressed in host cells and used
XX to develop diagnostic and therapeutic agents. Antagonists and
XX agonists based on the extracellular domains of P2U2 receptor, or
XX which affect receptor function by binding to one of the
XX intracellular domains, can be used to treat diseases caused by
XX aberrant activation of this receptor or to treat diseases whose
XX symptoms can be ameliorated by stimulating or inhibiting the
XX activity of the receptor.
XX
XX SQ Sequence 334 AA;
XX
XX Query Match 74.3%; Score 1226.5; DB 18; Length 334;
XX Best Local Similarity 71.5%; Pred. No. 5.8e-115;
XX Matches 226; Conservative 42; Mismatches 47; Indels 1; Gaps 1;
QY 1 MAONLSCENWLATEAILNKYVLSAFYATETIFGLLGNVTYVFGYLCMKWNSSNYLFN 60
Db 5 MAWNAATCKNWLAAEAPAELEKYLSIFYGIEFVGVLGNTIVVGYIFSLKWNSSNYLFN 64
QY 61 LSISDAFLCTPLILIKSYANDKGTGYDVLICISNRVYLVHTNLXTSLTLFTFISMDRYLLM 120
Db 65 LSVSDLAFLCTPLMIRSYANGNWIYGDVLCISNRVYLVHANLYTSILFTFISIDRYLLI 124
QY 121 KYPREHFLQKKEFAILISIAVWALVTLVPLMLTFINSVPKEEGSNCDIYASSGNPEHN 180
Db 125 KYPREHLLQKKEFAILISIAVWVLTLELLPLINPVITDNGTTCNDFASSGDPNPN 184
QY 181 LIYSICLTLGLFLPLSVCMCFYKVMVFLKRSQOQATALPLDKPQRLVLAUVIFSL 240
Db 185 LIYSWCLTLGLFLPLSVCMCFYKVMVFLKRSQOQATALPLDKPQRLVLAUVIFSVL 244
QY 241 FTPYHMRNRIASRLDSWPQ-GCTQKAIKSIYTLTRPLAFNLNSAINPIFYFLMGDHYRE 299
Db 245 FTPYHVMNRVRIASRLGSKWQYQCTQVINSFVITRPLAFNLNSVINPVFLLGDHFRD 304
QY 300 MLISKFRQYKSLTSF 315
Db 305 MLMNQLRHNFKSLTSF 320

```

Thu Apr 3 10:42:33 2003

65 LSVSDLAFLCTLPNLIRSYANGNIYGVLCISNRVVLHANLYTSILFTFISIDRYLII 124

121 KYPFHEFLOKKEFAILISLAVNALVTLEVLPMITFINSVPKEEGSNCDIYASSGNPEHN 180

125 KYPFHEHLLQKKEAILISLAWVLTLELLPILPINVTITDNGTTCNDFASSGDPNYN 184

181 LIYSCLTLLGLFPLSVCMCFYKMWVFLKRSQQOATALPLDKPQRLVWLAVWIFSL 240

185 LIYSCLTLLGLFPLSVCMCFYKMWVFLKRSQQOATALPLDKPQRLVWLAVWIFSVX 244

241 FTFPHMRNLRIASRLDSWPO-GCTOKAISKIYVTLTRPLAFNLNSAINPIFYFLMGDHYRE 299

245 FTFPHMRNLRIASRLDSWPO-GCTOKAISKIYVTLTRPLAFNLNSAINPIFYFLMGDHYRE 304

300 MLISKFRQYFKSLTSF 315

305 MLMQLRHNFKSLTSF 320

RESULT 8

AAU31029

ID AAU31029 standard; Protein; 387 AA.

XX AAU31029;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #1520.

XX Human; vaccination; gene therapy; nutritional supplement;

XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

XX vaccination, testing and therapy -

XX Claim 20; Page 392; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The

XX polypeptides and antibodies to the polypeptides are useful for

XX determining the presence of or predisposition to a disease associated

XX with altered levels of polypeptide. The polypeptides are also useful for

XX identifying agents (agonists and antagonists) that bind to them. Cells

XX expressing the proteins are useful for identifying a therapeutic agent

XX for use in treatment of a pathology related to aberrant expression or

XX physiological interactions of the polypeptide. Vectors comprising

XX the nucleic acids encoding the polypeptides and cells genetically

XX engineered to express them are also useful for producing the proteins.

XX The proteins are useful in genetic vaccination, testing and

XX therapy, and can be used as nutritional supplements. They may be used to

XX increase stem cell proliferation; to regulate haematopoiesis; and in

XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;

XX immune suppression and/or stimulation; as anti-inflammatory agents; and

XX in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid

XX sequences of novel human secreted proteins of the invention.

AAW22732

ID AAW22732 standard; Protein; 334 AA.

XX AAW22732;

XX 07-OCT-1997 (first entry)

XX Human ATP receptor.

XX ATP receptor; G-protein coupled receptor; agonist; antagonist.

XX Homo sapiens.

XX Key Location/Qualifiers

XX MISC-difference 212 /note= "encoded by TCC"

XX MISC-difference 235 /note= "encoded by TCG"

XX MISC-difference 244 /label= Unknown

XX /note= "encoded by CYT"

XX: WO9724929-A1.

XX 17-JUL-1997.

XX 11-JAN-1996; 96WO-US00392.

XX 11-JAN-1996; 96WO-US00392.

XX (HUMA-) HUMAN GENOME SCI INC.

XX LI Y;

XX WPI; 1997-372505/34.

XX N-PSDB; AAT75146.

XX Isolated human ATP receptor - agonists and antagonists of which are

XX useful in treatment of, e.g. asthma, hypertension, arterial

XX thrombosis and psychotic and neurological disorders

XX Claim 15; Fig 1A-C; 53pp; English.

XX Human ATP receptor (AAW22732) is structurally related to the G

XX protein-coupled receptor family. It shows 29.8% identity to a

XX murine P2u receptor. Its amino acid sequence was deduced from a

XX human placental cDNA clone (AAT75146). Recombinant ATP receptor can

XX be expressed in bacterial (e.g. E. coli), mammalian (e.g. COS) or

XX insect (e.g. Sf9) host cells and used to screen for agonists and

XX antagonists useful in the treatment of conditions related to

XX underexpression of the receptor (e.g. asthma, Parkinson's disease,

XX acute heart failure, hypotension, urinary retention and

XX osteoporosis) or overexpression of the receptor (e.g. arterial

XX thrombosis, hypertension, thrombolysis, angioplasty, cystic

XX fibrosis, ulcers, asthma, allergy, benign prostatic hypertrophy,

XX psychotic and neurological disorders, dyskinesias, endogenous

XX anorexia and bulimia).

XX Query Match

XX Best Local Similarity 73.9%; Score 1219.5; DB 18; Length 334;

XX Matches 225; Conservative 43; Mismatches 47; Indels 1; Gaps 1;

XX 1 MAQNLSNENLATEAILNKYILSAFYATEFTGLGNVTGVYGLFCMKNNSSNVYLEN 60

XX 5 MAWNATCNWLAEEAALBKLYLSIFGYIEFVGVGLNVTGVYGLFISLKNSSNLYLFN 64

XX 61 LSISDFALCTLPILIKSYANDKGTGVLCISNRVVLHANLYTSILFTFISIDRYLLM 120

SQ Sequence 387 AA:  
 Query Match 73.7%; Score 1216.5; DB 22; Length 387;  
 Best Local Similarity 71.1%; Pred. No. 7.2e-114;  
 Matches 224; Conservative 43; Mismatches 47; Indels 1; Gaps 1;  
 QY 2 AONLSCNLIATRAILNKKYLSAFYAEFFGLGNTVVFGYFCMKNNSSNYVLENL 61  
 DB 59 AWAATCKHLLAAEALEKYYLSIFGYEFGVGLNTIVVGYIFSLKNNSSNYVLENL 118  
 QY 62 SISDAFLCTPLILIKSYANDKTYGDVLCISNRYVHLNLTSLFTFTSMRDYLLMK 121  
 DB 119 SVSDLAFLCTPLMLRISSANWNIYGDVLCISNRYVHLNLTSLFTFTSIDRYLLIK 178  
 QY 122 YPREHFLOKKEFAILLISLAWALVTLEVLPMITFNSVPKREGSNCIDYASSGNPEHL 181  
 DB 179 YPREHLLQKKEFAILLISLAWVLTLELPLILINPVITDNGTFCNDFASSGDPNYNL 238  
 QY 182 IYSCLTLLGLFLPLSVCMCFYKVVFLKRRSQOQATLPLDKPQRLVLAIVTFSILF 241  
 DB 239 IYSCLTLLGLFSLPLFVCMCFYKVVFLKRRSQOQATLPLDKPQRLVLAIVTFSILF 298  
 QY 242 TPVHMRNLRIASRLDSWPQ-GCTQKAISYITLTPRPLAFINSAINPIFYFLMGDHYREM 300  
 DB 299 TPVHMRNRIASRLGSKQYQCTQVINSFYITRPLAFINSVINPVFYFLMGDHYREM 358  
 QY 301 LISKFRQYFKSLTSF 315  
 DB 359 LMNQLRHNFKSLTSF 373  
 RESULT 9  
 AAB45376  
 ID AAB45376 standard; Protein: 258 AA.  
 AC AAB45376;  
 XX 14-FEB-2001 (first entry)  
 DT Human secreted protein sequence encoded by gene 37 SEQ ID NO:128.  
 DE Human: secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 XX antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; vulnerable; gene therapy; autoimmune disease;  
 KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
 KW cerebrovascular disorder; nervous system disorder; infection; skin aging;  
 KW ocular disorder; wound healing; food additive; preservative.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000061628-A1.  
 XX  
 PD 19-OCT-2000.  
 PF 06-APR-2000; 2000WO-US09070.  
 XX  
 PR 09-APR-1999; 99US-0128695.  
 PR 14-JAN-2000; 2000US-0176052.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX  
 DR WPI; 2000-619228/59.  
 XX  
 PT New nucleic acid molecules encoding 49 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX  
 PS Disclosure; Page 447-448; 454pp; English.  
 XX  
 CC The polynucleotide sequences given in AAC81086 to AAC81134 encode the

CC human secreted proteins given in AAB45308 to AAB45356. AAB45357 to  
 CC AAB45384 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Examples of activities include:  
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;  
 CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;  
 CC neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 CC and vulnary. The polynucleotides and polypeptides can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
 CC disorders, cerebrovascular disorders, angiogenesis, nervous system  
 CC disorders, infections caused by bacteria, viruses and fungi and ocular  
 CC disorders. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
 CC components. AAC81077 to AAC81085 and AAB45307 represent sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 258 AA:  
 Query Match 61.5%; Score 1014.5; DB 21; Length 258;  
 Best Local Similarity 72.9%; Pred. No. 9e-94;  
 Matches 188; Conservative 33; Mismatches 36; Indels 1; Gaps 1;  
 QY 20 YVLSAFYAEFFGLGNTVVFGYFCMKNNSSNYVLENLSDFAFLCTPLILIKSY 79  
 DB 1 YVLSIFGYEFGVGLGNTIVVGYIFSLKNNSSNYVLENLSDFAFLCTPLILIRSY 60  
 QY 80 ANDKGTGYDVLICISNRYVHLNLTSLFTFTSMRDYLLMKYPFHEHLOKKEFALLIS 139  
 DB 61 ANGNLYGDVLCISNRYVHLNLTSLFTFTSIDRYLLIKYPFHEHLOKKEFALLIS 120  
 QY 140 LAWALVLEVLPMITFNSVPKREGSNCIDYASSGNPEHLNLTSLCTLLGLFLPLSV 199  
 DB 121 LAIVLVLTLELPLILINPVITDNGTFCNDFASSGDPNYNLISYMLCTLLGLFLPLF 180  
 QY 200 CFYKVVFLKRRSQOQATLPLDKPQRLVLAIVTFSILFTPYHMRNLRIASRLDSW 259  
 DB 181 CFYKVVFLKRRSQOQATLPLDKPQRLVLAIVTFSILFTPYHMRNLRIASRLDSW 240  
 QY 260 PQ-GCTQKAISYITL 276  
 DB 241 KQYQCTQVINSFYITR 258  
 RESULT 10  
 AAE04389  
 ID AAE04389 standard; Protein: 373 AA.  
 XX  
 AC AAE04389;  
 XX  
 DT 04-SEP-2001 (first entry)  
 XX  
 DE Human P2-purinergic receptor subtype, P2Y1.  
 KW Human: P2-purinergic receptor; P2Y1; cardiant; vasotropic; thrombolytic;  
 KW cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina;  
 KW myocardial infarction; ischaemic attack; preclampsia; bleeding disorder;  
 KW carotid endarterectomy; vascular graft surgery; brain disorder; migraine;  
 KW vascular injury; schizophrenia; eating disorder; depression; angioplasty;  
 KW peripheral vascular disease; platelet aggregation; restenosis; embolism;  
 KW thrombocytopaenic purpura; stroke; pertussis toxin-sensitive G protein;  
 KW Gi; disseminated intravascular coagulation; thrombosis.  
 XX  
 OS Homo sapiens.



XX SQ Sequence 373 AA; Query Match 28.7%; Score 474; DB 23; Length 373; Best Local Similarity 37.3%; Pred. No. 3.7e-39; Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;

QY 20 YLSAFYAIEFIFGLGNVTWVFGYLCMKNNSSNVYLFNLSISDFALCTLPILIKSY 79  
 DB 52 YLPAVYILVFIIGFLGNSVAIWMFVHMKPSGIVSVFMNLADFLYVLTLPALIFY 111  
 QY 80 ANDKG-TYGDVLCISNRVYVLTNLTSLFTFTISMDRYLLMKYPFEHFLQKKEFAILI 138  
 DB 112 FNKTDWIFGDAMCKLQRFIEHVNLYGSILFTLCISAHRYSGVYPLKSLGRLLKNAICI 171  
 QY 139 SLAVWALVLEVLPMLTFIN-VPKKEGSCNIDYASSGNPEHNLISLCLTLGLFLPLS 197  
 DB 172 SVLWLVIVVAISPILFSGTGVRKNKITCYDTSDEYLSYFIYSMCTTVAMFCVPLV 231  
 QY 198 VMCFYKMYVFLKRRSQOQATLPL-DKQRLVLAIVLAVVIFSLFTFYHMRNLRIASRL 256  
 DB 232 LILGCYGLIVRALIYKDLDNS---PLRRKSIYLVITVFAVSYPFFHVMKTMNLRARL 288  
 QY 257 D-SWPGQCT-QKAISKIYTLTRPLAFINSAINPIFYFLMGDHYREML 301  
 DB 289 DFQTPAMCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTERRRL 335

RESULT 12  
 AAU10984  
 ID AAU10984 standard; Protein; 373 AA.  
 XX AAU10984;  
 XX  
 DT 12-MAR-2002 (first entry)  
 DE Purinergic receptor P2Y, G-protein coupled 1, isoform #1.  
 KW Purinergic receptor P2Y, G-protein coupled 1; P2RY1; anticoagulant;  
 KW coagulant; platelet aggregation; haplotyping; drug screening;  
 KW transgenic animal; human.  
 XX Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc-difference 34 /note= "Wild type Ala substituted by Val"  
 FT  
 XX WO200190117-A2.  
 XX 29-NOV-2001.  
 XX 21-MAY-2001; 2001WO-US16432.  
 XX 19-MAY-2000; 2000US-205996P.  
 XX (GENA-) GENAISSANCE PHARM INC.  
 XX Kazemi A, Koshy B, Tanguay DA;  
 XX WPI; 2002-083074/11.  
 DR  
 XX New purinergic receptor P2Y G-protein coupled 1 (P2RY1) gene  
 PT polymorphic variants, useful e.g. in studying the expression and  
 PT function of P2RY1 and screening candidate drugs for treating diseases  
 PT related to P2RY1 activity  
 XX  
 XX Disclosure; Page -; 79pp; English.  
 PS  
 CC The invention relates to a novel isolated polypeptide comprising a  
 CC sequence which is a polymorphic variant of a reference sequence for the  
 CC purinergic receptor P2Y, G-protein coupled, 1 (P2RY1) protein or its  
 CC fragment. The polymorphic variant comprises one or more variant amino

CC acids selected from valine at a position 34 and glycine at a position 262. The polymorphic variants are useful in studying the expression and function of P2RY1, in expressing P2RY1 protein for use in screening for candidate drugs to treat diseases related to P2RY1 activity, in studying the effect of the variation on the biological activity of P2RY1, and the binding affinity of candidate drugs targeting P2RY1 for the treatment of disorders related to platelet aggregation. The haplotyping methods are useful in validating P2RY1 as a candidate target for treating a specific condition or disease predicted to be associated with P2RY1 activity, or in the design of clinical trials of candidate drugs for treating a specific condition or disease associated with P2RY1 activity. The transgenic animals are useful for studying expression of the P2RY1 isogenes in vivo, for in vivo screening and testing of drugs targeted against P2RY1 protein, and for testing the efficacy of therapeutic agents and compounds for disorders related to platelet aggregation in a biological system. The present sequence represents the amino acid sequence of human purinergic receptor P2Y, G-coupled protein 1 (P2RY1), isoform #1.  
 CC Note: This sequence is not shown in the specification but is derived from the human wild type P2RY1 sequence given in Figure 3 (see AAU10983).  
 CC  
 XX SQ Sequence 373 AA;  
 Query Match 28.7%; Score 474; DB 23; Length 373; Best Local Similarity 37.3%; Pred. No. 3.7e-39; Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;

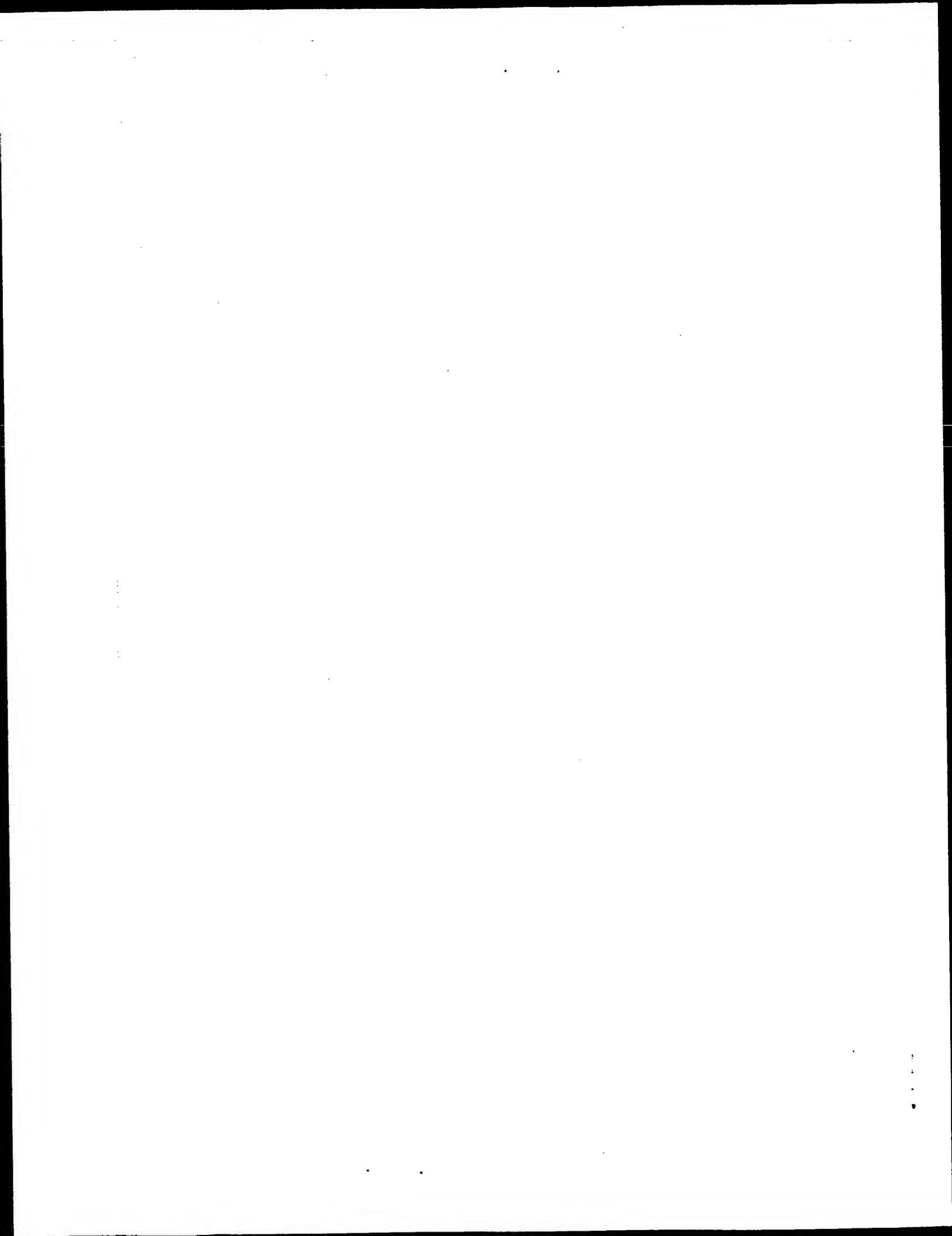
QY 20 YLSAFYAIEFIFGLGNVTWVFGYLCMKNNSSNVYLFNLSISDFALCTLPILIKSY 79  
 DB 52 YLPAVYILVFIIGFLGNSVAIWMFVHMKPSGIVSVFMNLADFLYVLTLPALIFY 111  
 QY 80 ANDKG-TYGDVLCISNRVYVLTNLTSLFTFTISMDRYLLMKYPFEHFLQKKEFAILI 138  
 DB 112 FNKTDWIFGDAMCKLQRFIEHVNLYGSILFTLCISAHRYSGVYPLKSLGRLLKNAICI 171  
 QY 139 SLAVWALVLEVLPMLTFIN-VPKKEGSCNIDYASSGNPEHNLISLCLTLGLFLPLS 197  
 DB 172 SVLWLVIVVAISPILFSGTGVRKNKITCYDTSDEYLSYFIYSMCTTVAMFCVPLV 231  
 QY 198 VMCFYKMYVFLKRRSQOQATLPL-DKQRLVLAIVLAVVIFSLFTFYHMRNLRIASRL 256  
 DB 232 LILGCYGLIVRALIYKDLDNS---PLRRKSIYLVITVFAVSYPFFHVMKTMNLRARL 288  
 QY 257 D-SWPGQCT-QKAISKIYTLTRPLAFINSAINPIFYFLMGDHYREML 301  
 DB 289 DFQTPAMCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTERRRL 335

RESULT 13  
 AAU10985  
 ID AAU10985 standard; Protein; 373 AA.  
 XX AAU10985;  
 XX  
 DT 12-MAR-2002 (first entry)  
 DE Purinergic receptor P2Y, G-protein coupled 1, isoform #2.  
 KW Purinergic receptor P2Y, G-protein coupled 1; P2RY1; anticoagulant;  
 KW coagulant; platelet aggregation; haplotyping; drug screening;  
 KW transgenic animal; human.  
 XX Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc-difference 262 /note= "Wild type Val substituted by Gly"  
 FT  
 XX WO200190117-A2.  
 XX 29-NOV-2001.  
 XX









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 15:13:52 ; Search time 19 Seconds  
(without alignments)  
490.898 Million cell updates/sec

Title: US-09-891-138A-2  
Perfect score: 1650  
Sequence: 1 MAQNLSCENWLATEAILNKY.....REMLTSKPROVFKSLTSPT 317

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pap: \*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pap: \*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pap: \*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pap: \*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pap: \*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1226.5	74.3	334	2	US-08-559-524A-2
2	1226.5	74.3	334	3	US-08-749-707-2
3	474	28.7	373	2	US-08-559-524A-4
4	474	28.7	373	3	US-08-749-707-4
5	473.5	28.7	362	3	US-08-513-974B-374
6	370.5	22.5	373	3	US-08-513-974B-373
7	353	21.4	355	1	US-08-153-848-28
8	353	21.4	355	3	US-09-299-843A-28
9	353	21.4	355	3	US-09-299-843A-32
10	353	21.4	355	3	US-09-088-337B-28
11	353	21.4	355	4	US-09-088-337B-32
12	353	21.4	355	5	PCT-US93-11153-32
13	353	21.4	355	5	PCT-US93-11153-32
14	353	21.4	355	5	PCT-US93-11153-32
15	338	20.5	328	3	US-08-513-974B-39
16	338	20.5	328	3	US-08-513-974B-371
17	336.5	20.4	360	4	US-08-875-573-20
18	336.5	20.4	360	4	US-09-232-878-2
19	336.5	20.4	360	4	US-09-045-583-55
20	336.5	20.4	360	4	US-09-534-185-55
21	335.5	20.3	375	1	US-08-442-134A-2
22	335.5	20.3	375	1	US-08-444-581B-2
23	335.5	20.3	375	1	US-08-446-088A-2
24	335.5	20.3	375	2	US-08-559-524A-3
25	335.5	20.3	375	3	US-08-749-707-3
26	329	19.9	355	4	US-09-045-583-53
27	329	19.9	355	4	US-09-534-185-53

28 328 19.9 355 1 US-08-012-988A-2 Sequence 2, Appli  
29 328 19.9 355 1 US-08-450-393A-5 Sequence 5, Appli  
30 328 19.9 355 4 US-08-446-669-5 Sequence 5, Appli  
31 328 19.9 355 4 US-09-239-938-1 Sequence 1, Appli  
32 328 19.9 355 5 PCT-US95-00476-5 Sequence 5, Appli  
33 326.5 19.8 360 4 US-08-833-752-10 Sequence 10, Appli  
34 325 19.7 302 2 US-08-467-948A-30 Sequence 30, Appli  
35 325 19.7 302 3 US-08-467-947A-30 Sequence 30, Appli  
36 325 19.7 328 3 US-08-513-974B-56 Sequence 36, Appli  
37 325 19.7 328 3 US-08-513-974B-380 Sequence 380, App  
38 324.5 19.7 327 3 US-08-513-974B-372 Sequence 372, App  
39 322 19.5 328 3 US-08-459-046-2 Sequence 2, Appli  
40 316.5 19.2 384 3 US-09-071-434-3 Sequence 3, Appli  
41 315 19.1 355 1 US-08-461-244-2 Sequence 2, Appli  
42 315 19.1 355 4 US-09-045-583-56 Sequence 56, Appli  
43 315 19.1 355 4 US-09-534-185-56 Sequence 56, Appli  
44 314 19.0 344 2 US-08-467-948A-8 Sequence 8, Appli  
45 314 19.0 344 3 US-08-467-947A-8 Sequence 8, Appli

## ALIGNMENTS

RESULT 1  
US-08-559-524A-2  
: Sequence 2, Application US/08559524A  
: Patent No. 5871963  
: GENERAL INFORMATION:  
: APPLICANT: Conley, Pamela B.  
: APPLICANT: Jantzen, Haus-Michael  
: TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
: NUMBER OF SEQUENCES: 14  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
: STREET: 1800 M Street, N.W.  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: USA  
: ZIP: 20036-5869  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA: US/08/559,524A  
: FILING DATE: 15-NOV-1995  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Adler, Reid G.  
: REGISTRATION NUMBER: 30,988  
: REFERENCE/DOCKET NUMBER: 044481-5010-00-US  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 202-467-7000  
: TELEFAX: 202-467-7176  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 334 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-559-524A-2

Query Match 74.3% Score 1226.5; DB 2; Length 334;  
Best Local Similarity 71.5%; Pred. No. 3.5e-93;  
Matches 226; Conservative 42; Mismatches 47; Indels 1; Gaps 1;  
QY 1 MAQNLSCENWLATEAILNKYLSAFYAEIFGLLGNVTVFYGLFCMKNNSSNYLFN 60  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 5 MAWNACTKNWLAEEAALEKYYLSIFYGIEFVGVGLNVTIVVYGYIFSLKNWSSNYLFN 64  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 LSISDFAFCLTPIILIKSYANDKGTGYDVLICISNRYVHFNLTYSILFTFTISMORYLIM 120  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 125 KYPFRHLLQKKEFAILLISLAIWLVLTLELLPILPINVITDNGTNCDFASSGDPNTN 184  
QY 181 LIYSCLTLLGLFPLSVCMCFEYKVMVFLKRRSQOQATALPLDKPQLRVLAVVIFSL 240  
Db 185 LIYSMCLTLLGLFPLFVCMCFEYKVMVFLKRRSQOQATALPLDKPQLRVLAVVIFSVL 244  
QY 241 FTPYHIMRNLRASRLDSWPO-GCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYRE 299  
Db 245 FTPYHVMNRVIRASRLGSKWQYQCTQVINSFYITRALGFLNSVINPVFYFLMGDHYRE 304  
QY 300 MLISKFRQYFKSLTSF 315  
Db 305 MLMNQLRHNFKSLTSF 320

RESULT 3  
US-08-559-524A-4  
; Sequence 4, Application US/08559524A  
; Patent No. 5871963  
; GENERAL INFORMATION:  
; APPLICANT: Conley, Pamela B.  
; APPLICANT: Jantzen, Hans-Michael  
; TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
; STREET: 1800 M Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5869  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/559,524A  
; FILING DATE: 15-NOV-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adler, Reid G.  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 044481-5010-00-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-467-7000  
; TELEFAX: 202-467-7176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 373 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-559-524A-4  
Query Match 28.7%; Score 474; DB 2; Length 373;  
Best Local Similarity 36.6%; Pred. No. 1.6e-31;  
Matches 105; Conservative 58; Mismatches 116; Indels 8; Gaps 6;

QY 20 YLLSAFYAIEFIFGLGNVTGVVFGYLFCKMKNSSNVYLFNLSISDFALCTLPILIKSY 79  
Db 52 YLPAVAVILVFIIGFLGNSVAIWMFVHMKPWSGISVYMFNLALADFLVLTLPALIFY 111  
QY 80 ANDKG-TYGDVLCISNRYVLTNLTSLFIFISMDRYLLMKYPRHFLQKKEFAILL 138  
Db 112 FNKTDFGDMCKLQRFIFHVNLYGSLFLTCISAHRYSGVVYPLKSLGLKKNNAVYI 171  
QY 139 SLAWALVTLEVLPMLTFFINS-VPKKEGNSNCIDYASSGNPEHNLITSLCTLLGLFLPLS 197  
Db 172 SVLWLVVVGISPIIFYSGTIGIRKNTITCYDTTSDEYLSRYFYISMCCTVAMFCVPLV 231  
QY 198 VMCFYKVMVFLKRRSQOQATALPL-DKPQLRVLAVVIFSLFPTPYHIMRNLRASRL 256

Db 65 LSVSLAFLCTLPMLIRSYANGNWIYGVLCISNRYVLANLYTSILFTFISIDRYLII 124  
QY 121 KYPFRHLLQKKEFAILLISLAIWLVLTLELLPILPINVITDNGTNCDFASSGDPNTN 184  
Db 125 KYPFRHLLQKKEFAILLISLAIWLVLTLELLPILPINVITDNGTNCDFASSGDPNTN 184  
QY 181 LIYSCLTLLGLFPLSVCMCFEYKVMVFLKRRSQOQATALPLDKPQLRVLAVVIFSL 240  
Db 185 LIYSMCLTLLGLFPLFVCMCFEYKVMVFLKRRSQOQATALPLDKPQLRVLAVVIFSVL 244  
QY 241 FTPYHIMRNLRASRLDSWPO-GCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYRE 299  
Db 245 FTPYHVMNRVIRASRLGSKWQYQCTQVINSFYITRALGFLNSVINPVFYFLMGDHYRE 304  
QY 300 MLISKFRQYFKSLTSF 315  
Db 305 MLMNQLRHNFKSLTSF 320

RESULT 2  
US-08-749-707-2  
; Sequence 2, Application US/08749707  
; Patent No. 6063582  
; GENERAL INFORMATION:  
; APPLICANT: Conley, Pamela B.  
; APPLICANT: Jantzen, Hans-Michael  
; TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
; STREET: 1800 M Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5869  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/749,707  
; FILING DATE: 15-NOV-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adler, Reid G.  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 044481-5010-01-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-467-7000  
; TELEFAX: 202-467-7176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 334 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-749-707-2  
Query Match 74.3%; Score 1226.5; DB 3; Length 334;  
Best Local Similarity 71.5%; Pred. No. 3.5e-93;  
Matches 226; Conservative 42; Mismatches 47; Indels 1; Gaps 1;

QY 1 MAQNLSCEWLATEAILNKYLSAFYAEIFGLGNVTGVVFGYLFCKMKNSSNVYLFN 60  
Db 5 MAWNATCKNWLAAEALKEKYLISFYEGFEVGVGLNITVVGYIFSLKKNSSNIVLFN 64  
QY 61 LSISDFALCTLPILIKSYANDKGTYGDVLCISNRYVLTNLTSLFIFISMDRYLLM 120  
Db 65 LSVSLAFLCTLPMLIRSYANGNWIYGVLCISNRYVLANLYTSILFTFISIDRYLII 124  
QY 121 KYPFRHLLQKKEFAILLISLAIWLVLTLELLPILPINVITDNGTNCDFASSGDPNTN 184

Db 232 LILGCGYGLIVRALIYKDLNDS---PLRKSIIYLIIVLTFAVSIYIPFHWKTMNLRRL 288  
 QY 257 D-SWPGCT-QKATKSIYTLTRPLAFNSAINPIFYLMGDHYREML 301  
 Db 289 DFQTPMCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 4

US-08-749-707-4  
 ; Sequence 4, Application US/08749707  
 ; Patent No. 6063582  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Conley, Pamela B.  
 ; APPLICANT: Jantzen, Hans-Michael  
 ; TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
 ; STREET: 1800 M Street, N.W.  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20036-5869  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/749,707  
 ; FILING DATE: 15-NOV-1996  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Adler, Reid G.  
 ; REGISTRATION NUMBER: 30,988  
 ; REFERENCE/DOCKET NUMBER: 044481-5010-01-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-467-7000  
 ; TELEFAX: 202-467-7176  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 373 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-749-707-4

Query Match 28.7%; Score 474; DB 3; Length 373;  
 Best Local Similarity 36.6%; Pred. No. 1.6e-31;  
 Matches 105; Conservative 58; Mismatches 116; Indels 8; Gaps 6;  
 QY 20 YLSAFATIEFTGLGNVTVVGYLFCMKNNSSNVYLFNLISIDEAFGLCTLPILIKSY 79  
 Db 52 YVLPVAVYILVFTIGFLGNSVAWMEVFHMKPSGISVYMFNLADFLYVUTFLPALIYY 111  
 QY 80 ANDKG-TYGDVLCISNRVYHNLVTSILFTFTISMDRYLLMKYFPREHFQKKEFALLI 138  
 Db 112 FNKTWIFGDAMCKLQRFIFHNVLSILFTCTISAHRYSGVYVPLKSLGRLLKKNVYI 171  
 QY 139 SLAVVALVTVLEVLPLMTFINS-VPEEGSCNIDYASSNPENHNLISYCLLLGLFLPLS 197  
 Db 172 SVLWLVVVGISPLIFSTGGIRKNKTTTCYDTSDEYLSRYSIYSCITVAMEFCVPLV 231  
 QY 198 VMCFYYKVMVFLKRRSQOQATAPL-DKQRLVVLAVVIFSLFTPVHIMNRLRIASRL 256  
 Db 232 LILGCGYGLIVRALIYKDLNDS---PLRKSIIYLIIVLTFAVSIYIPFHWKTMNLRRL 288  
 QY 257 D-SWPGCT-QKATKSIYTLTRPLAFNSAINPIFYLMGDHYREML 301  
 Db 289 DFQTPMCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 5

US-08-513-974B-374  
 ; Sequence 374, Application US/08513974B  
 ; Patent No. 6114139  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hinuma, Shuji  
 ; APPLICANT: Hosoya, Masaki  
 ; APPLICANT: Fujii, Ryo  
 ; APPLICANT: Ohtaki, Tetsuya  
 ; APPLICANT: Fukusumi, Shoji  
 ; APPLICANT: Ohgi, Kazuhiro  
 ; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
 ; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
 ; NUMBER OF SEQUENCES: 380  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/513,974B  
 ; FILING DATE: 14-SEP-1995  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/JF95/01599  
 ; FILING DATE: 10-AUG-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 7-093989  
 ; FILING DATE: 19-AUG-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 7-057186  
 ; FILING DATE: 16-MAR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 7-007177  
 ; FILING DATE: 20-JAN-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-326611  
 ; FILING DATE: 28-DEC-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-270017  
 ; FILING DATE: 02-NOV-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-236357  
 ; FILING DATE: 30-SEP-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-236356  
 ; FILING DATE: 30-SEP-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-189274  
 ; FILING DATE: 11-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-189273  
 ; FILING DATE: 11-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 6-189272  
 ; FILING DATE: 11-AUG-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Resnick, David S.  
 ; REGISTRATION NUMBER: 34,235  
 ; REFERENCE/DOCKET NUMBER: 45753  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-523-3400  
 ; TELEFAX: 617-523-6440  
 ; INFORMATION FOR SEQ ID NO: 374:  
 ; SEQUENCE CHARACTERISTICS:

Thu Apr 3 10:42:34 2003

us-09-891-138a-2.ra1

```

;
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-513-974B-374

Query Match      28.7%; Score 473.5; DB 3; Length 362;
Best Local Similarity 35.9%; Pred. No. 1.7e-31;
Matches 110; Conservative 57; Mismatches 122; Indels 17; Gaps 8;

QY 10 WLA-----TEAILNK-----YILSAFYALIEFIFGLLGNVTTFVGYLFCMKNNSSNVYLFN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 WAAGNATTCKSLTKTGQFYLLPTVYILVITGFLGNSVAIWMFVHMRPNWSGISVYMFN 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 LSTDSFAFLCTLPILIKSYANDKG-TYGDVLCISNRYVLTNLYSILFLFISMDRYLL 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 LALADELYVLTLPALIFYFNKTDWIFGDVCKLQRFIFHVNLYSGILFLTCISVHRVYTG 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 120 MKYPFREFHLOKKEFALLISLAWALVTLEVLPMLTFTIN-SVPKEEGSNCIDYASSGNPE 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 VVHPLKSLGRKKNAVYSSSLWALVAVIAPIIFYSGIGYVRNKTITCYDTTADAYLR 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 179 HNLISLCLTLGLFLPLSLMCFYKMYVFLKRRSQOQATLPL-DKQRLVLAUVIF 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 SYEVSMCTVFECFPIFVILGCGYGLIVKALYKDLNDS---PLRRKSTYLVIIIVLTVF 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 238 SILFTYHYHMRNRIASRLD-SWPQCT-QKAIKSYITLTRPLATLNSAINPIFYFLMGD 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 AVSYLPFHVMKTLNLRALDFTQPMCAFNDKYATYQVTRGLASLNSCVDPILYFLAGD 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 296 HYREML 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 TFRRL 324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-08-513-974B-373
; Sequence 373, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 45753
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 373:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-513-974B-373

Query Match      22.5%; Score 370.5; DB 3; Length 373;
Best Local Similarity 33.5%; Pred. No. 4.6e-23;
Matches 94; Conservative 50; Mismatches 130; Indels 7; Gaps 5;

QY 19 KY-YLSAFYAIEFIFGLLGNVTTFVGYLFCMKNNSSNVYLFNLSISDFALCTLPILIK 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 KYVLLPVSVGVVGLGCLNVVALYIFLCRLKINWASTYMFHLAVSDLSYAAASLPLLVY 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 SYA-NDKGTGYDVLCSNRYVLTNLYSILFLFISMDRYLLMKYPPREHFLQKKEFAI 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 YYAGDHPFFSTVLCVKLVRFLLTYNLKYSILFLTCISVHRCLGVLRPLHSLRWGRARYAR 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 137 LISLAWALVTLEVLPMLTFTIN-SVPKEEGSNCIDYASSGNPEHNLISLCTLLGLFLPL 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 RVAAVVWLVLACQAPVLYFVTTTSVRGTRITCHDTSARELFSHFVAYSSVLMGLLFAVPF 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 197 SYMCFEYKMYVFLKRRSQOQATLPLDKPQ--RLVLAUVIFSLFTPYHYHMRNRIAS 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 SVILVCYVLMARRLLKPAYGTGGLPRAKRSVTRIALVLAVALCFLPFHVRTLYYSF 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 255 RLDSMPQCG-TOKAIKSYITLTRPLAFLNSAINPIFYFLMG 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 R-SLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAG 310
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-08-153-848-28
; Sequence 28, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
```



```

; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweiart, Vicki L.
; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,848
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5759804and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-153-848-28

```

```

Query Match 21.4%; Score 353; DB 1; Length 355;
Best Local Similarity 28.1%; Pred. No. 1.2e-21;
Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYVLSAFYAEIFGLGNVTYVFGYLFCKMKNSSNYLFLNLSISDFACLTLPIL 75
Db 28 VFGTVFLSIFYVIFAIGLVGNLLVFAALTNSKKPKSVTDIYLLNLSLDLFLVATLPFW 87
Qy 76 IKSANDKGTGVDVLCISNRVYVLTNLTSTLFTFISMDRYLLMKYPFREFHLOKKEFA 135
Db 88 THYLINKEGLH-NAMCKFTTAFFFGFGSIFFIIVISIDRYLAIVLAANSMMNRVQHG 146
Qy 136 ILISLAVNALVTLVPLMTFTNSVPKEGSGNCI-DYASSGNPEHNLISLCLTLGLFLI 194
Db 147 VTISLGVMAAAILVAAPQFMF---TKQENECGLDYPVLOETWPLRVNVTNFGFL 202
Qy 195 PLSNVCFYFKMW-VFLKRSQOQATALPDKPQRLVAVVIFSIETPTPHIMRLRIA 253
Db 203 PLLIMSYCFRIQTFLFCKNKKAKAI-----KLILLVIVFFLFTWPTYNVMIETL 256
Qy 254 SRLDSWPGCTOKAISKIYTLTRPLAFNSAINPIFFYFLMGDHYREMLISKFRQYFSL 312
Db 257 KLYDFPSCDMRKDLRLALSVTETVAFSHCCNLPLIYAFAGEKFRRYL---YHLYGKCL 312

```

```

RESULT 8
US-08-153-848-32
; Sequence 32, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweiart, Vicki L.

```

```

; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,848
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5759804and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-153-848-32

```

```

Query Match 21.4%; Score 353; DB 1; Length 355;
Best Local Similarity 28.1%; Pred. No. 1.2e-21;
Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYVLSAFYAEIFGLGNVTYVFGYLFCKMKNSSNYLFLNLSISDFACLTLPIL 75
Db 28 VFGTVFLSIFYVIFAIGLVGNLLVFAALTNSKKPKSVTDIYLLNLSLDLFLVATLPFW 87
Qy 76 IKSANDKGTGVDVLCISNRVYVLTNLTSTLFTFISMDRYLLMKYPFREFHLOKKEFA 135
Db 88 THYLINKEGLH-NAMCKFTTAFFFGFGSIFFIIVISIDRYLAIVLAANSMMNRVQHG 146
Qy 136 ILISLAVNALVTLVPLMTFTNSVPKEGSGNCI-DYASSGNPEHNLISLCLTLGLFLI 194
Db 147 VTISLGVMAAAILVAAPQFMF---TKQENECGLDYPVLOETWPLRVNVTNFGFL 202
Qy 195 PLSNVCFYFKMW-VFLKRSQOQATALPDKPQRLVAVVIFSIETPTPHIMRLRIA 253
Db 203 PLLIMSYCFRIQTFLFCKNKKAKAI-----KLILLVIVFFLFTWPTYNVMIETL 256
Qy 254 SRLDSWPGCTOKAISKIYTLTRPLAFNSAINPIFFYFLMGDHYREMLISKFRQYFSL 312
Db 257 KLYDFPSCDMRKDLRLALSVTETVAFSHCCNLPLIYAFAGEKFRRYL---YHLYGKCL 312

```

```

RESULT 9
US-09-299-843A-28
; Sequence 28, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweiart, Vicki L.
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66

```

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,843A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/088,337  
; FILING DATE: 01-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/153,848  
; FILING DATE: 17-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,452  
; FILING DATE: 17-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jill E. Uhl  
; REGISTRATION NUMBER: 43,213  
; REFERENCE/DOCKET NUMBER: 27866/32059B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 355 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-299-843A-28

Query Match 21.4%; Score 353; DB 3; Length 355;  
Best Local Similarity 28.1%; Pred. No. 1.2e-21;  
Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;  
QY 16 ILNKYLSAFAIEFIFGLGNVTVFGYLCMKNNSSNVYLFNLSISDFAPLCTLPIL 75  
Db 28 VFGTVLSIFYSVIFAIGLVGNLLVFPALTNSKKPKSVTDIYLLNLSLDLFLVATLPFW 87  
QY 76 IKSYANDKGTGYGVLCISNRYVLTNLTYSILFTFISMDRYLLMKYPFREFHFLQKKEFA 135  
Db 88 THYLINEKGLH-NAMCKETFAFFFGIFGIFFTVISIDRYLAIVLAANSMNRTVOHG 146  
QY 136 ILISLAWALVTLEVPMLTFINSVPKEGNCI-DYASSGNPEHNLISLCTLLGLFLI 194  
Db 147 VTISLGWAAAILVAAPQPMF---TKOKENECGDYPEVLQEIWPVLRNVTNFGFL 202  
QY 195 PLSVMCFYKMW-VFLKRRSQOQATALPLDKPQRLVAVLAVIFSILTPYHIMNRLIA 253  
Db 203 PLLIMSYCFRIQTFLSCNHHKAKAI-----KLILLVIVFVFWTPYVNMIFLET 256  
QY 254 SRLDSWPGCGTQAKISYITLTPRLAFLNSAINPIFFELMGDHYREMLISKFRQYFKSL 312  
Db 257 KLYDFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAFAGEKFRYL---YHLYGKCL 312

RESULT 10  
US-09-299-843A-32  
; Sequence 32, Application US/09299843A  
; Patent No. 6107475  
; GENERAL INFORMATION:  
; APPLICANT: Godiska, Ronald

; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,843A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/088,337  
; FILING DATE: 01-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/153,848  
; FILING DATE: 17-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,452  
; FILING DATE: 17-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jill E. Uhl  
; REGISTRATION NUMBER: 43,213  
; REFERENCE/DOCKET NUMBER: 27866/32059B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 355 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-299-843A-32

Query Match 21.4%; Score 353; DB 3; Length 355;  
Best Local Similarity 28.1%; Pred. No. 1.2e-21;  
Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;  
QY 16 ILNKYLSAFAIEFIFGLGNVTVFGYLCMKNNSSNVYLFNLSISDFAPLCTLPIL 75  
Db 28 VFGTVLSIFYSVIFAIGLVGNLLVFPALTNSKKPKSVTDIYLLNLSLDLFLVATLPFW 87  
QY 76 IKSYANDKGTGYGVLCISNRYVLTNLTYSILFTFISMDRYLLMKYPFREFHFLQKKEFA 135  
Db 88 THYLINEKGLH-NAMCKETFAFFFGIFGIFFTVISIDRYLAIVLAANSMNRTVOHG 146  
QY 136 ILISLAWALVTLEVPMLTFINSVPKEGNCI-DYASSGNPEHNLISLCTLLGLFLI 194  
Db 147 VTISLGWAAAILVAAPQPMF---TKOKENECGDYPEVLQEIWPVLRNVTNFGFL 202  
QY 195 PLSVMCFYKMW-VFLKRRSQOQATALPLDKPQRLVAVLAVIFSILTPYHIMNRLIA 253  
Db 203 PLLIMSYCFRIQTFLSCNHHKAKAI-----KLILLVIVFVFWTPYVNMIFLET 256  
QY 254 SRLDSWPGCGTQAKISYITLTPRLAFLNSAINPIFFELMGDHYREMLISKFRQYFKSL 312  
Db 257 KLYDFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAFAGEKFRYL---YHLYGKCL 312

RESULT 11  
US-09-088-337B-28



Thu Apr 3 10:42:34 2003

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

[illegible]

RESULT 14  
PCT-US93-11153-32  
: sequence 32, Application PC/TUS9311153

```

GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweikart, Vicki L.
TITLE OF INVENTION: Novel Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11153
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-11153-32

```

```

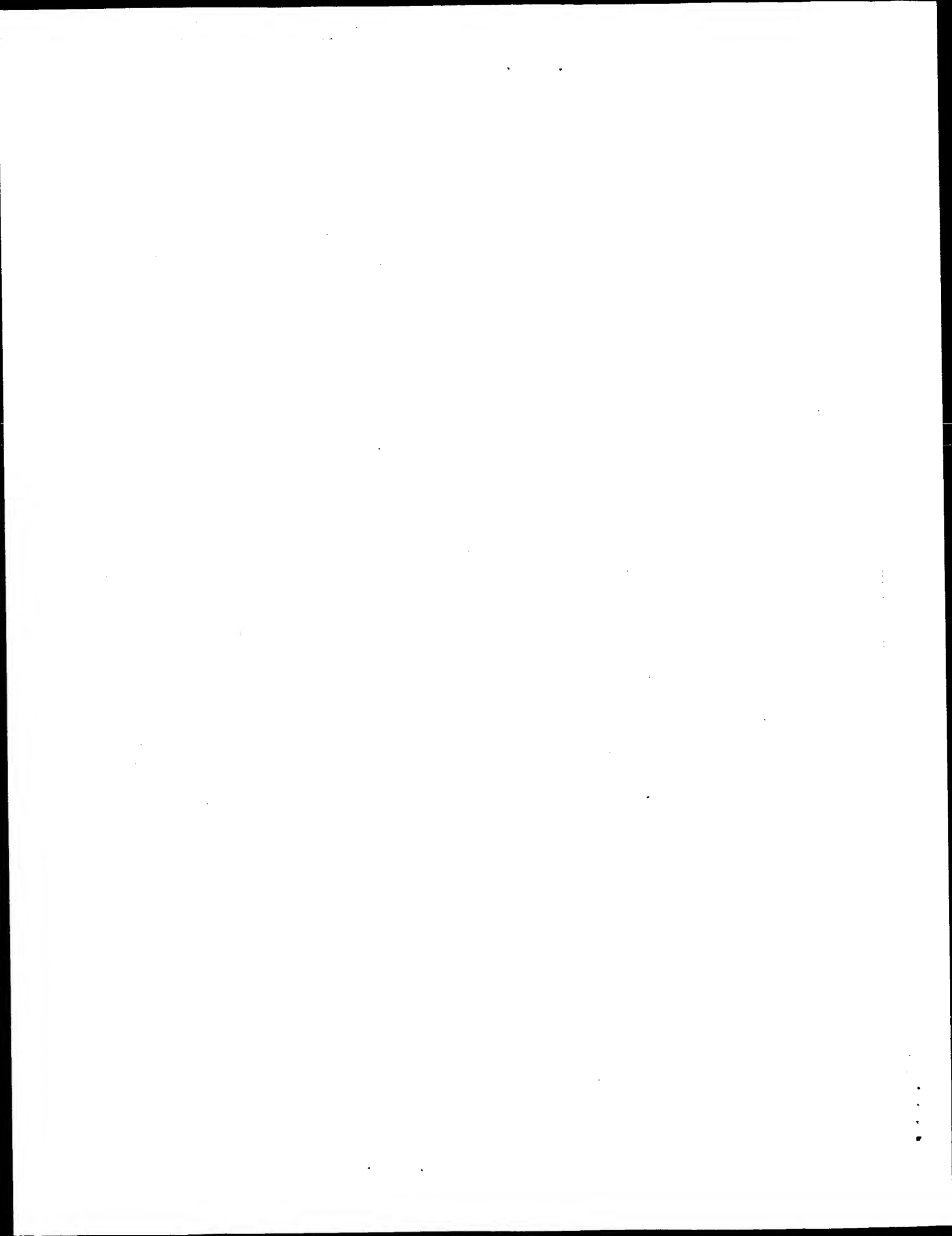
Query Match      21.4%; Score 353; DB 5; Length 355;
Best Local Similarity 28.1%; Score 1.1.2e-21;
Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

16 ILNKYILSAFAIEFIEGLGNVTYVFGYLCFMKNWNSNYLVNLSIDSDFAELCTPIL 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
28 VFGIVFLSIFYSVIFAIGLVGNLLVVFALTNSKPKSVTDYLLNLALSDLLEVFATPEW 87
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
76 IKSXANDKGTGYDGVLCISNRYVLHTNLTYLSILETFISMDRYLLMKYPFHEHFKQKEFA 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
88 THYLINERGLH-NAMCKRFTTAFPFEGFGSIFPTVISIDRYLAIVLAANSNNETVQHG 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 ILISLVWALVTLEVLPMLFINSVPKKEGSNCI-DYASSGNEPHNLVYSLCTLGLGFLI 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147 VTISLGVWAAAILVAAPQFHF---TKQKENECLGDYPEVLQEIWPLRVNVTNFGFL 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
195 PLVSMCFYYKMW-VFLKRRSQOATALPDKQRLWVLAVYFESTLETPYHMMNLRIA 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203 PLLINSYCYFRITQITLFSCNHHKAKAI-----KULLVWVFFLFTWYVNMVLFETL 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
254 SRLDSWPGCGTQALKSIYTLTRPLAFNLAINPFIYFLMGDHYRMLLSKFERQYFKSL 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 KLYDFEPPSDMKRDKURLALSVEITVAFSHCCINPLIYAFAGEFRRYL--VHLYGKCL 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-08-513-974B-39
; Sequence 39, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 15:17:12 ; Search time 23 Seconds  
(without alignments)  
842.614 Million cell updates/sec

Title: us-09-891-138a-2

Perfect score: 1650

Sequence: 1 MAONLSCENWLATEAILNKY.....REMLSKPQYFKSLTSPT 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1231.5	74.6	334	10	US-09-765-034-2	Sequence 2, Appli
2	1224.5	74.2	334	9	US-10-270-587-2	Sequence 2, Appli
3	479	29.0	373	9	US-10-092-135-7	Sequence 7, Appli
4	475	28.8	299	9	US-10-270-144-4	Sequence 4, Appli
5	474	28.7	373	9	US-10-092-135-5	Sequence 5, Appli
6	474	28.7	373	9	US-10-092-135-5	Sequence 6, Appli
7	472.5	28.6	362	9	US-10-092-135-3	Sequence 3, Appli
8	472.5	28.6	362	9	US-10-092-135-4	Sequence 4, Appli
9	426	25.8	337	9	US-10-023-7758-2	Sequence 2, Appli
10	426	25.8	337	9	US-10-270-144-2	Sequence 2, Appli
11	426	25.8	337	10	US-09-943-798-4	Sequence 4, Appli
12	373.5	22.6	341	9	US-10-270-587-3	Sequence 4, Appli
13	366	22.2	276	10	US-09-943-798-2	Sequence 2, Appli
14	353	21.4	355	10	US-09-789-482-4	Sequence 4, Appli
15	353	21.4	355	10	US-09-789-486-4	Sequence 4, Appli
16	336.5	20.4	360	9	US-10-120-394-20	Sequence 20, Appli
17	336.5	20.4	360	9	US-09-764-413-20	Sequence 20, Appli
18	336.5	20.4	360	10	US-09-837-446-2	Sequence 2, Appli
19	336.5	20.4	360	10	US-09-796-744-17	Sequence 17, Appli

20	328	19.9	355	10	US-09-961-068-1	Sequence 1, Appli
21	328	19.9	355	10	US-09-960-547-1	Sequence 1, Appli
22	326.5	19.8	330	10	US-09-826-791-2	Sequence 2, Appli
23	326.5	19.8	346	9	US-09-828-478-2	Sequence 2, Appli
24	326.5	19.8	346	10	US-09-826-791-6	Sequence 6, Appli
25	326.5	19.8	346	10	US-09-866-230-7	Sequence 7, Appli
26	326.5	19.8	346	10	US-09-866-230-9	Sequence 9, Appli
27	326.5	19.8	360	10	US-09-938-719-10	Sequence 10, Appli
28	326.5	19.8	360	10	US-09-938-719-10	Sequence 10, Appli
29	326.5	19.8	360	10	US-09-938-703-10	Sequence 10, Appli
30	325	19.7	302	9	US-10-024-494-30	Sequence 30, Appli
31	315	19.1	355	10	US-09-104-752-2	Sequence 2, Appli
32	314.5	19.1	363	10	US-09-864-761-42849	Sequence 42849, A
33	314	19.0	344	9	US-10-024-494-8	Sequence 8, Appli
34	309.5	18.8	339	9	US-09-828-478-4	Sequence 4, Appli
35	309.5	18.8	339	10	US-09-848-889-12	Sequence 12, Appli
36	309.5	18.8	339	9	US-09-788-133-2	Sequence 2, Appli
37	309.5	18.8	367	9	US-09-828-478-6	Sequence 6, Appli
38	309	18.7	355	10	US-09-938-719-9	Sequence 9, Appli
39	309	18.7	355	10	US-09-939-226-9	Sequence 9, Appli
40	309	18.7	355	10	US-09-938-703-9	Sequence 9, Appli
41	307.5	18.6	369	10	US-09-823-114-9	Sequence 9, Appli
42	307	18.6	362	9	US-09-940-240-20	Sequence 20, Appli
43	303	18.4	391	9	US-09-990-940-20	Sequence 20, Appli
44	302.5	18.3	348	10	US-09-827-937A-17	Sequence 17, Appli
45	301	18.2	332	9	US-10-001-835-140	Sequence 140, App

## ALIGNMENTS

RESULT 1  
US-09-765-034-2  
; Sequence 2, Application US/09765034  
; Patent No. US20020137887A1  
; GENERAL INFORMATION:  
; APPLICANT: Hedrick, Joseph A.  
; APPLICANT: Lachowicz, Jean E.  
; APPLICANT: Wang, Wei  
; APPLICANT: Gustafson, Eric L.  
; TITLE OF INVENTION: Adenosine Receptor  
; FILE REFERENCE: CN01084  
; CURRENT APPLICATION NUMBER: US/09765,034  
; CURRENT FILING DATE: 2001-01-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO. 2  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-765-034-2

Query Match	74.6%	Score 1231.5;	DB 10;	Length 334;
Best Local Similarity	71.8%	Pred. No. 4.1e-101;		
Matches 227;	Conservative 42;	Mismatches 46;	Indels 1;	Gaps 1;
OY	1	MAONLSCENWLATEAILNKYLSAFYALEFFGLGNVTVVGVYFCMKWNSSNYLFFN 60		
Db	5	MAWNAATCKNWLAAEALEKYYLSIFYGIEFVGVGLNTIVVGYIFSLKWNSSNYLFFN 64		
OY	61	LSISDFALCTLPILIKSYANDKGTGDLVLCISNRYVHLNLTSLIFLFFISMDRYLLM 120		
Db	65	LSVSDLAFLCTLPMLIRSYANGNIYGVLCISNRYVHLNLTSLIFLFFISIDRYLLI 124		
OY	121	KYPFHEFLQKKEFAILLISLAWALVTLEVLPMLETFINSVPKEGSCNIDYASSGNPEHN 180		
Db	125	KYPFHEFLQKKEFAILLISLAWALVTLEVLPMLETFINSVPKEGSCNIDYASSGNPEHN 184		
OY	181	LIYSLCTLLGLFLPLSVCMCFYKVMVFLKRSQOQATALPDKPQRLWLAWLWIFSTL 240		
Db	185	LIYSLCTLLGLFLPLSVCMCFYKVMVFLKRSQOQATALPDKPQRLWLAWLWIFSTL 244		
OY	241	FTPHIMNLRBIASRLDSWFO-GCTQKAIKSIYTLTRFLAFLNSAINPIFYFLMGDHYRE 299		







```

; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-775B-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNVYLFNLSISDFAFCTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVVIITYIFKMPKSSIIIMLNACTDLLYLTSLPFLI 90
QY 77 KSYAN-DKGYGDVLCISNRYVLTNLYTSILFTFISMDRYLLMKYPPREHFLQKKEA 135
Db 91 HYASGENWIFGDFMCKFIHFSEFHNLYSSILFTCFIPRYCVIIHPMSPCSFIHKTRCA 150
QY 136 ILISLAWVALVTLVPLMTFINSVPKEEGSNCIDYASSGNPEHNL--YSLCLTLLGL 193
Db 151 VVACAVVYIISLVAVIPMTFLTITSTNTRNSACLDLTSS--DELNIKWNLLITATTC 208
QY 194 IPLSVMCFFYYKMYVFLKRRSQOQATALPLDKPQRLVVLAVWIFSIILFTPYHMRNLRIA 253
Db 209 LPLVIVILCYTIIHTLTHGLQTDSC--KQARRLTILLALLAFYVGFPHILRVIRIE 266
QY 254 SRLDSWPGQCT-QKAISYIYTLRPLAFINSAINPIFYFLMGDHYREMLISKFR 306
Db 267 SRLLS--ISCSIEQIHEAYIVSRPLAALTFTGNLLLYVVVSNFQAVGVSTVR 318

RESULT 10
US-10-270-144-2
; Sequence 2, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
; US-10-270-144-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNVYLFNLSISDFAFCTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVVIITYIFKMPKSSIIIMLNACTDLLYLTSLPFLI 90

; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 362
; TYPE: PRT
; ORGANISM: MELEAGRIS GALLOPAVO
; US-10-092-135-4

Query Match 28.6%; Score 472.5; DB 9; Length 362;
Best Local Similarity 35.9%; Pred. No. 3.8e-34;
Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;

QY 10 WLA---TEAILNK-----YLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNVYLFN 60
Db 22 WAAGNASTKCSLTKTGQFYIPLTVIYITGFLGNSVAIMWFHMPWGSISVYXEN 81
QY 61 LSISDFAFCTLPILIKSYANDKG--TYGDVLCISNRYVLTNLYTSILFTFISMDRYLL 119
Db 82 LALADFLYVLTLPALLIFYFNKTDWIFGDVCMCKLQRFIEHVNLYGSILFTCISVHRXTG 141
QY 120 MKYPPREHFLQKKEFALLISLAWVALVTLVPLMTFINS--VPKEEGSNCIDYASSGNPE 178
Db 142 VVHPLKSLGRKKKNVYVSSLVWALVAVIAPILFYSGTVRRNKITTCYDTTADAYLR 201
QY 179 HNLISYSLCTLLGLFLIPLSVCMCFYKMYVFLKRRSQOQATALPL--DKPQRLVVLAVVIF 237
Db 202 SYFVSMCTTVFMECFPIFVILGCGYGLIKVLYKLDNS---PLRKSIIYLVIIIVTVF 258
QY 238 SILFTPYHMRNLRIASRLD-SWPQGCT-QKAISYIYTLRPLAFINSAINPIFYFLMGD 295
Db 259 AVSVLPFHVMKTLNLRARLDFQTPOMCAFNDKVYATQVTRGLASLNSCVDPILYFLAGD 318
QY 296 HYREML 301
Db 319 TFRRL 324

RESULT 9
US-10-023-775B-2
; Sequence 2, Application US/10023775B
; Publication No. US2003002282A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (EP/GB) only
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)
; APPLICANT: Fidock, Mark David
; TITLE OF INVENTION: No. US2003002282A1e1 Polypeptide
```

```

; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-775B-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNVYLFNLSISDFAFCTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVVIITYIFKMPKSSIIIMLNACTDLLYLTSLPFLI 90
QY 77 KSYAN-DKGYGDVLCISNRYVLTNLYTSILFTFISMDRYLLMKYPPREHFLQKKEA 135
Db 91 HYASGENWIFGDFMCKFIHFSEFHNLYSSILFTCFIPRYCVIIHPMSPCSFIHKTRCA 150
QY 136 ILISLAWVALVTLVPLMTFINSVPKEEGSNCIDYASSGNPEHNL--YSLCLTLLGL 193
Db 151 VVACAVVYIISLVAVIPMTFLTITSTNTRNSACLDLTSS--DELNIKWNLLITATTC 208
QY 194 IPLSVMCFFYYKMYVFLKRRSQOQATALPLDKPQRLVVLAVWIFSIILFTPYHMRNLRIA 253
Db 209 LPLVIVILCYTIIHTLTHGLQTDSC--KQARRLTILLALLAFYVGFPHILRVIRIE 266
QY 254 SRLDSWPGQCT-QKAISYIYTLRPLAFINSAINPIFYFLMGDHYREMLISKFR 306
Db 267 SRLLS--ISCSIEQIHEAYIVSRPLAALTFTGNLLLYVVVSNFQAVGVSTVR 318

RESULT 10
US-10-270-144-2
; Sequence 2, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
; US-10-270-144-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNVYLFNLSISDFAFCTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVVIITYIFKMPKSSIIIMLNACTDLLYLTSLPFLI 90
```

QY	48	MKNWSSNYVLFNLISDSFAFLCTLPILIKSYAN-DKGYGDVLCISNRYVLTNLTYSI	106
Db	1	MRPKSSTIIMLNACTDLYLTSLPFLHYASGENWIFGDPMKCFIRFSPHFNLYSSI	60
QY	107	LFTFTSMORYLMMKYPRFHFQKKEFAILISLAYWALVTLEVLPMLTFFINSVPKEGS	166
Db	61	LFTCFISIRYCVIIHPMCSFSLHKTRCAVACAVWILISLAVIPMTLIITSTNETNRS	120
QY	167	NCIDYASSGNPENLI--YSLCTLTGLFTPLVSMGCFYKWKVFLKRRSQOQATALPJD	224
Db	121	ACLDLTSS--DELNTIKWNLILTATFCLPLVIVTLCYTTIIHTLTHGLQTDSCLL--KQ	176
QY	225	KPQRLVVLAVIFSILFTPYHIMRNIRIARSLDWSPPQGT--QKAISKSYITLTPRLAFINS	283

Db 177 KARRLTILLLLAFYVCFPLFPFHILRVIRIESRLLS---ISCSIENTQIHEAVIYSRPLAALNT 234

Qy 284 AINPIFYFLMGDHYREMLSKFR 306

Db 235 FGNLLLYVVVSDNFQAVCVSTVR 257

## RESULT 14

US-09-789-482-4

; Sequence 4, Application US/09789482

; Patent No. US20020054875A1

; GENERAL INFORMATION:

; APPLICANT: Koch, Alisa E.

; APPLICANT: Ruth, Jeffrey H.

; APPLICANT: Rotman, James B.

; TITLE OF INVENTION: Therapeutic Methods That Target

; FILE OF INVENTION: Fractalkine or CX3CR1

; FILE REFERENCE: 3238.1000-003

; CURRENT APPLICATION NUMBER: US/09/789,482

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: US 60/183,568

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-789-482-4

Query Match 21.4%; Score 353; DB 10; Length 355;

Best Local Similarity 28.1%; Pred. No. 1.3e-23;

Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYILSAFYAIEFIFGLLGNVTYVFGYLFCKMKNWSSNVYLFNLSISDFALCTLPIL 75

Db 28 VFGTVFLSIFYVIFAIGLVGNLTVFALTNSKKPKSVTDIYLLNLALSDDLFLVATLPFW 87

Qy 76 IKSYANDKGYGDVLCISNRYVLTNLTYSILFTFISMDRYLLMKYPPEHFLQKKEFA 135

Db 88 THYLINEKGLH-NAMCKFTTAFFIFGFGSIFITVISIDRYLAIVLAANSNMNRVQHG 146

Qy 136 ILISLAWALVTLEVLPMLTFFINSVPKEGNCI-DYASSGNPEHNLIIYSLCTLGLFLI 194

Db 147 VTISLGWAAAILVAAPQPMF-----TKQENECIGDYPEVLQEIWPVLRNVETNFGFL 202

Qy 195 PLSVMCFYKMW-VFLKRSQOQATPLDKPQRLVLAIVFISILFTPYHIMRNLRIA 253

Db 203 PLLIMSYCFRIIQTFLSCNHHKAKAI-----KLILLVIVFFLFTWTPYVNMIFLET 256

Qy 254 SRLDSWPQGCQKAISYITLTPFLNLSAINPIFYFLMGDHYREMLISKFRQYFKSL 312

Db 257 KLYDFFPSCDMRKDLRLALSVTETVAFSHCHCLNPLIYAFAGEKFRYL---YHLYGKCL 312

## RESULT 15

US-09-789-486-4

; Sequence 4, Application US/09789486

; Patent No. US20020055456A1

; GENERAL INFORMATION:

; APPLICANT: Koch, Alisa E.

; TITLE OF INVENTION: Therapeutic Methods That Target

; FILE OF INVENTION: Fractalkine or CX3CR1

; FILE REFERENCE: 3238.1000-004

; CURRENT APPLICATION NUMBER: US/09/789,486

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: US 60/183,568

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-789-486-4

Query Match 21.4%; Score 353; DB 10; Length 355;

Best Local Similarity 28.1%; Pred. No. 1.3e-23;

Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYILSAFYAIEFIFGLLGNVTYVFGYLFCKMKNWSSNVYLFNLSISDFALCTLPIL 75

Db 28 VFGTVFLSIFYVIFAIGLVGNLTVFALTNSKKPKSVTDIYLLNLALSDDLFLVATLPFW 87

Qy 76 IKSYANDKGYGDVLCISNRYVLTNLTYSILFTFISMDRYLLMKYPPEHFLQKKEFA 135

Db 88 THYLINEKGLH-NAMCKFTTAFFIFGFGSIFITVISIDRYLAIVLAANSNMNRVQHG 146

Qy 136 ILISLAWALVTLEVLPMLTFFINSVPKEGNCI-DYASSGNPEHNLIIYSLCTLGLFLI 194

Db 147 VTISLGWAAAILVAAPQPMF-----TKQENECIGDYPEVLQEIWPVLRNVETNFGFL 202

Qy 195 PLSVMCFYKMW-VFLKRSQOQATPLDKPQRLVLAIVFISILFTPYHIMRNLRIA 253

Db 203 PLLIMSYCFRIIQTFLSCNHHKAKAI-----KLILLVIVFFLFTWTPYVNMIFLET 256

Qy 254 SRLDSWPQGCQKAISYITLTPFLNLSAINPIFYFLMGDHYREMLISKFRQYFKSL 312

Db 257 KLYDFFPSCDMRKDLRLALSVTETVAFSHCHCLNPLIYAFAGEKFRYL---YHLYGKCL 312

Search completed: April 2, 2003, 15:19:49

Job time : 23 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	474	28.7	373	2	JC4162	P2Y receptor - bo
2	474	28.7	373	2	JC4737	G protein-coupled
3	473.5	28.6	362	2	S33733	G protein-coupled
4	473.5	22.6	373	2	A47556	ATP receptor P2u
5	353	21.4	355	2	JC4304	orphan G protein-c
6	352	21.3	365	2	S68679	G protein-coupled
7	336.5	20.4	360	2	A57160	chemokine (C-C) re
8	336	20.4	328	2	I55450	G protein-coupled
9	335.5	20.3	375	2	A54946	G protein-coupled
10	333	20.2	308	2	I50241	P-2U nucleotide re
11	332.5	20.2	354	2	I58186	G protein-coupled
12	332	20.1	328	2	JC4800	probable G protein
13	328	19.9	355	2	A45177	P2Y6 receptor - hu
14	326.5	19.8	360	2	A54587	chemokine (C-C) re
15	325.5	19.7	388	2	JN0605	chemokine (C-C) re
16	318.5	19.3	359	2	A48921	somatostatin recep
17	317.5	19.2	355	2	A55733	interleukin-8 rece
18	316.5	19.2	384	2	A47249	G protein-coupled
19	315.5	19.1	355	2	J01231	brain-specific som
20	315	19.1	355	2	JC5067	interleukin-8 rece
21	314	19.0	369	2	JC2083	G protein-coupled
22	314	19.0	369	2	B41795	somatostatin recep
23	314	19.0	370	2	JC5549	somatostatin recep
24	313	19.0	344	2	T09508	heptahelical P2Y5-
25	312.5	18.9	360	2	A53611	intron 17 purinerg
26	311	18.8	358	2	A53752	interleukin-8 rece
27	309.5	18.8	384	2	JC4629	interleukin-8 rece
28	309	18.7	355	2	I49339	somatostatin recep
29	308	18.7	369	2	D41795	macrophage inflamm
						somatostatin recep

[illegible]

RESULT 2

JC4737  
 G protein-coupled receptor P2Y1 - human  
 N:Alternate names: P2Y1 purinergic receptor; P2Y1 purinoreceptor  
 C:Species: Homo sapiens (man)  
 C:Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 17-Nov-2000  
 C:Accession: JC4737; J04615; S54253  
 R:Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M.  
 Biochem. Biophys. Res. Commun. 211, 588-593, 1996  
 A:Title: Cloning and tissue distribution of the human P2Y1 receptor.  
 A:Reference number: JC4737; MUID:96205320; PMID:8630005  
 A:Accession: JC4737  
 A:Molecule type: DNA  
 A:Residues: 1-373 <JAN>  
 A:Cross-references: GB:S81950; NID:g1839438; PIDN:AA847091.1; PID:g1839439  
 R:Ayyanathan, K.; Webbs, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.P.  
 Biochem. Biophys. Res. Commun. 218, 783-788, 1996  
 A:Title: Cloning and chromosomal localization of the human P2Y1 purinoreceptor.  
 A:Reference number: JC4615; MUID:96158962; PMID:8579591  
 A:Accession: JC4615  
 A:Molecule type: mRNA  
 A:Residues: 1-373 <AYY>  
 A:Cross-references: GB:U04209; NID:g1147730; PIDN:AAA97872.1; PID:g1147731  
 A:Experimental source: erythro leukemia cells  
 R:Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Cloning of a human putative P2Y receptor.  
 A:Reference number: S54253  
 A:Accession: S54253  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-137, 139-373 <LEO>  
 A:Cross-references: EMBL:249205; NID:g798835; PIDN:CAA9066.1; PID:g798836  
 C:Comment: This receptor belongs to a family of G protein-coupled receptors. It responds  
 C:Genetics:  
 A:Gene: P2Y1; GDB:P2Y1  
 A:Cross-references: GDB:677125; OMIM:601167  
 A:Map position: 3pter-3qter  
 C:Superfamily: ATP receptor P2u  
 C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein  
 F:52-77/Domain: transmembrane #status predicted <TM1>  
 F:88-111/Domain: transmembrane #status predicted <TM2>  
 F:124-152/Domain: transmembrane #status predicted <TM3>  
 F:171-191/Domain: transmembrane #status predicted <TM4>  
 F:214-237/Domain: transmembrane #status predicted <TM5>  
 F:261-282/Domain: transmembrane #status predicted <TM6>  
 F:305-328/Domain: transmembrane #status predicted <TM7>  
 F:11,27,113,197/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted  
 F:330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depen

Query Match 28.7%; Score 474; DB 2; Length 373;  
 Best Local Similarity 37.3%; Pred. No. 4.2e-32;  
 Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;  
 QY 20 YLSAFYAIEFIFGLGNVTWVFYGLFCMKNNSSNVFLNLSISDFAFELCTLPILIKSY 79  
 Db 52 YLPNAVILVFIFGLNSVAIMWFVFMKPSGSIYVFNALADFLVLTLPALIFVY 111  
 QY 80 ANDKG-TYGDVLCISNRYLHTNLYTSILFTFISMDRYLLMKYPPREHFLQKFAILI 138  
 Db 112 FNKTDWIFGDMCKLQRFIFHVNLYSILFTICISAHRYSGVYVPLKSLGRKKNAICI 171  
 QY 139 SLAWVALVTLVLEPMLTFLINS-VPKESGNCIDYASSGNFHNLYSLCTLLGLFLPLS 197  
 Db 172 SVLVWLVVAISPILFYSYSGTVGRNKNITTCYDTSDEYLRSLRFSYISMCYTVAMFCVPLV 231  
 QY 198 VMCFYFKMVFVFLKRRSQOQATALPL-DKQRLVVLAVVIFSIPTFYHMRNLRTASRL 256  
 Db 232 LILGCYGLIVRALIYKDLNDS--PLRKSISYLVIIIVLTVFVAVSYIPFHVMTNLRARL 288

QY 257 D-SWQOQCT-QKAISKIYTLTRPLAFNLNSAINPIFYFLMGDHYREML 301  
 Db 289 DFOTPAWCAFNDRYATYQVTRGLASLNSCVDPIYFLAGDTFRRL 335

RESULT 3

S33733  
 G protein-coupled receptor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-Sep-1999  
 C:Accession: S33733  
 R:Webb, T.E.; Simon, J.; Krishek, B.J.; Bateson, A.N.; Smart, T.G.; King, B.F.; Burns  
 FEBS Lett. 324, 219-225, 1993  
 A:Title: Cloning and functional expression of a brain G-protein-coupled ATP receptor.  
 A:Reference number: S33733; MUID:93285340; PMID:8508924  
 A:Accession: S33733  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-362 <WEB>  
 A:Cross-references: EMBL:X73268; NID:g395084; PIDN:CAA51716.1; PID:g395085  
 C:Superfamily: ATP receptor P2u  
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 28.6%; Score 472.5; DB 2; Length 362;  
 Best Local Similarity 35.9%; Pred. No. 5.4e-32;  
 Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;  
 QY 10 WLA----TEAILNK-----YLSAFYAIEFIFGLGNVTWVFYGLFCMKNNSSNVYLEN 60  
 Db 22 WAGNATTKCSLTGTFQFYVLTPTVITGFLGNSVAIMWFVFMKPSGSIYVFN 81  
 QY 61 LSISDFAFELCTLPILIKSYANDKG-TYGDVLCISNRYLHTNLYTSILFTFISMDRYLL 119  
 Db 82 LALADEFLVLTLPALIFYFNKTDWIFGDMCKLQRFIFHVNLYSILFTICISVHRYTG 141  
 QY 120 MKYPERHFLOKKEFAILISLAWVALVTLVLEPMLTFLINS-VPKESGNCIDYASSGNPE 178  
 Db 142 VVHLKSLGRLLKKNVAVVSSLVWALVAVIAPILFYSGTGVRNKNITTCYDTSDEYLR 201  
 QY 179 HNLYSLCTLLGLPLSLVSMCFYFKMVFVFLKRRSQOQATALPL-DKQRLVVLAVVIF 237  
 Db 202 SYFVYSMTCTVFMEFCIPIFIVLGCYGLIVKALIIYKDLNDS---PLRKSISYLVIIIVTFV 258  
 QY 238 SILTPYHINLRTASRLD-SWQOQCT-QKAISKIYTLTRPLAFNLNSAINPIFYFLMGD 295  
 Db 259 AVSYLPFHVMTNLRARLDPQTQMCACAFNDKVYATYQVTRGLASLNSCVDPIYFLAGD 318  
 QY 296 HYREML 301  
 Db 319 TFRRL 324

RESULT 4  
 A47556  
 ATP receptor P2u - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1999  
 C:Accession: A47556  
 R:Lustig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993  
 A:Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.  
 A:Reference number: A47556; MUID:93281707; PMID:7685114  
 A:Accession: A47556  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-373 <LUS>  
 A:Cross-references: GB:L14751; NID:g309457; PIDN:AAA39871.1; PID:g309458  
 C:Superfamily: ATP receptor P2u  
 C:Keywords: transmembrane protein

Query Match 22.6%; Score 373.5; DB 2; Length 373;  
 Best Local Similarity 33.5%; Pred. No. 1e-23;



Query Match	21.4%	Score 353;	DB 2;	Length 355;
Best Local Similarity	28.1%;	Pred. No. 5e-22;		
Matches	84;	Conservative 66;	Mismatches 133;	Indels 16; Gaps 6;
QY	16	ILNKYLSAFYAIEIFIGLLGNVTWFGYLCFKMNNNSVYLFNLISIDEAFLCTLPIL	75	
Db	28	VFGTVFSIFTSVIFAIGLVGNLTVVALNSKKPKSVTDIYLLNLAUSDLLFVATLPFW	87	
QY	76	IKSYANDKGTGYDVLCTISNRVYLVHTNLYTSILFLTFISMDRYLLMKYPFHEFLOKKEFA	135	
Db	88	THYLNKGLH-NAMCKETTAFFIGFPGSIFITVISIDRYLAIVLAANSNNRTVQHG	146	
QY	136	ILISLAVNALVLEVLPLMTFINSVPKEGSNCT-DYASSGNPEHNLVSYSLCTLLGLLI	194	
Db	147	VTISLGVMAAAILVAAPQEMF-----TQKENECLGDYPEVLQEIWPVLRVNETNLFGLL	202	
QY	195	PLSVNCFYFYKMW-VFLKRRSQOQATALPLDKQRLVLAIVFISLTFPYHIMNLRIA	253	
Db	203	PILLINSYCFRIOTLFSCKNHAKAI-----KILWVAVTFVLTFTVNVMTLETL	256	

A:Gene: GDB:CMKBR4  
A:Cross-references: GDB:677463  
A:Map position: 3p21-3p21  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein  
F:40-65/Domain: transmembrane #status predicted <TM1>

F:76-97/Domain: transmembrane #status predicted <TM2>  
F:112-133/Domain: transmembrane #status predicted <TM3>  
F:151-175/Domain: transmembrane #status predicted <TM4>  
F:208-226/Domain: transmembrane #status predicted <TM5>  
F:243-264/Domain: transmembrane #status predicted <TM6>  
F:291-308/Domain: transmembrane #status predicted <TM7>  
F:29-276,110-187/disulfide bonds: #status predicted  
F:72,350/Binding site: phosphate (Ser) (covalent) (by protein kinase II) #status predicted  
F:145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted  
F:183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 20.4%; Score 336.5; DB 2; Length 360;  
Best Local Similarity 29.4%; Pred. No. 1.2e-20;  
Matches 93; Conservative 60; Mismatches 132; Indels 31; Gaps 9;

QY 14 EATLNKYL-----SAF-----YAEIFGGLGNVTGVFGYLFCKMKNW 52  
DB 13 ESIYNYIYESTPKCTKEGKAFGELEPLPLYSLVFVGLGNSVVVL-VLFYKRLR 71  
QY 53 S-SNVYLFNLSISDAFLCTPLILKSYANDKGTGVDVLCISNRVYVLTNLTSLFTLF 111  
DB 72 SMTDVLNLAISDLLEFVSPW-CYAAADQWVFGGLCKMISWMYLVGFYSIGIFFVWL 130  
QY- 112 ISMDRYLLMKYPPEHFLQKKEFAILLISLAVVALVTLVPLMLTFINSVPKEGNCIDY 171  
DB 131 MSIDRYLAIHVAFSLRARTLTGVITSLATSVAVFSLPGLFSTCTYERNHYCKTK 190  
QY 172 ASSGNPEHNLISLCLTLGLFPLSVMGFFYKVMVFLKRSQQOQATLPLDKFORLV 231  
DB 191 YLSNSTTWKVLSSLEINILGLVPLGIM-LFCYSMII-----RTLQHCNKKNAVKMIF 245  
QY 232 LAVVIFSLFTPYHIMRNLIARLSWPGQCT-QRAIKSIYTLTRPLAFLNSAINPIFY 290  
DB 246 AVVFLGFTPTNYVILFETLLEVL-QDCTFERLYDAIQTATLFAVHCCUNPIIY 304  
QY 291 FLMGDHVREMLISKFR 306  
DB 305 FFLGEEKFRKYLQFK 320

RESULT 8  
I55450  
G protein-coupled P2 receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 19-May-2000  
C:Accession: I55450  
R:Chang, K.; Hanaoka, K.; Kumada, M.; Takuwa, Y.  
J. Biol. Chem. 270, 26152-26158, 1995  
A:Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.  
A:Reference number: I55450; MUID:96064682; PMID:7592819  
A:Accession: I55450  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <RES>  
A:Cross-references: GB:D63665; NID:g1066007; PIDN:BA009816.1; PID:g1066008  
A:Superfamily: ATP receptor P2u  
C:Keywords: G protein-coupled receptor

Query Match 20.4%; Score 336.5; DB 2; Length 328;  
Best Local Similarity 29.9%; Pred. No. 1.2e-20;  
Matches 85; Conservative 52; Mismatches 137; Indels 10; Gaps 4;

QY 22 LSFAFAIEFGLGNVTGVFGYLFCKMKNWSSNVYLFNLSISDAFLCTPLILKSYA- 80  
DB 29 LPPVSVVVLVGLPLNVCVIAQICASRTLTRSAVYVLTNLTALADLLIYACSLPLIYNAR 88  
QY 81 NDKGTGVDVLCISNRVYVLTNLTSLFTPLISMDRYLLMKYPPE-HFLQKKEFAILLIS 139  
DB 89 GDHWFGDLACRLVRFYANLHGSILFTLCISFORVILGICHLAPLAPWKRGRRAAANVVC 148  
QY 140 LAVVALVTLVPLMLTFINSVPKEGNCIDYASSGNPEHNLISLCLTLGLFPLISVM 199

Db 149 GVWLVVTAQCLTAVFAATGIQNRNRTCYDLSPPLSTRYLPGMALVIGFLPDLPTAL 208  
QY 200 CFYYKMWVFLKRSQQOQATLPL-----DKPORLVVLAIVFSLFTPYHIMRNLIAR 254  
Db 209 LACYCRMA---RRLCRODGPAGPAOERSKAARMAVVAEVISFLPHITKTAYLAV 265  
QY 255 RLDSWPQCGTQKAIKSIYTLTRPLAFLNSAINPIFYFLMGDHYR 298  
Db 266 RSPGVSCPVLETFAAAYKGRFPASANSVLDPLIFVFTQOKFR 309  
RESULT 9  
A54946  
P-2U nucleotide receptor - human  
C:Species: Homo sapiens (man)  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 17-Mar-1999  
C:Accession: A54946  
R:Par, C.E.; Sullivan, D.M.; Paradiso, A.M.; Lazarowski, E.R.; Burch, L.H.; Olsen, J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 3275-3279, 1994  
A:Title: Cloning and expression of a human P-2U nucleotide receptor, a target for cys  
A:Reference number: A54946; MUID:94211846; PMID:8159738  
A:Accession: A54946  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-375 <PAR>  
A:Cross-references: GB:U07225  
A:Note: parts of this sequence were confirmed by protein sequencing  
C:Genetics:  
A:Gene: GDB:P2RY2; HP20; P2U  
A:Cross-references: GDB:362713; OMIM:600041  
A:Map position: 11q13.5-11q14.1  
C:Superfamily: ATP receptor P2u  
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 20.3%; Score 335.5; DB 2; Length 375;  
Best Local Similarity 33.1%; Pred. No. 1.5e-20;  
Matches 93; Conservative 52; Mismatches 127; Indels 9; Gaps 7;

QY 19 KY-YLSAFYAIEFTFGLGNVTGVFGYLFCKMKNWSSNVYLFNLSISDAFLCTPLILK 77  
Db 32 KYVLLPVSVGVVGLGCLNAGVLYIFLCRLKTNASTYMFHLAVSDALYAAASLPLVY 91  
QY 78 SYA-NDKGTGVDVLCISNRVYVLTNLTSLFTPLISMDRYLLMKYPPEHFLQKKEFAI 136  
Db 92 YYARGDHPFSTVLCIKLVRFELFYTNLYCISILFICISVHRCLGVLRPLSLRWGRARYAR 151  
QY 137 LISLAWVALVTLVPLMLTFINSVPKEGNCIDYASSGNPEHNLISLCLTLGLFLIPL 196  
Db 152 RVAGAVVVLVLAQAPVLYFVTTTSARGP-LTCHDTSAPELFSRFVAYSSVMGLLFAVFF 210  
QY 197 SVMCFYKMWVFLKRSQQOQATLPLDKPORLVVLAIVFSLFTPYHIMRNLIAR 254  
Db 211 AVILVCYVLMARRLLKPAYTSGGLPRAKRSKSVRTIAVLAVALFCLFPFHVTRTLIYXF 270  
QY 255 RLDSWPQCGTQKAIKSIYTLTRPLAFLNSAINPIFYFLMG 294  
Db 271 R-SLDLSCHTLNAINMAYKVT-R-LASANSCLDPLVLYFLAG 308

RESULT 10  
I50241  
G protein-coupled receptor 6H1 - chicken  
N:Alternate names: purinoceptor 6H1  
C:Species: Gallus gallus (chicken)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Jun-2000  
C:Accession: I50241; JC4618  
R:Kaplan, M.H.; Smith, D.I.; Sundick, R.S.  
J. Immunol. 151, 628-636, 1993  
A:Title: Identification of a G protein coupled receptor induced in activated T cells.  
A:Reference number: I50241; MUID:93329058; PMID:8393036  
A:Accession: I50241  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

Query Match	20.2%;	Score 332.5;	DB 2;	Length 354;
Best Local Similarity	27.9%;	Pred. No. 2.6e-20;		

Db 149 VAVWLAVTTQCLPTAFAATGIORNTVCYDLSPPALATHYMPYGMALTVIGFLPPFAAL 208  
QY 200 CFYYKVMVFLKRRSQOQATAPL-----DKPQLVLLAVVIFSLFTPYHIMRNLRAS 254  
Db 209 LACYCLLAC---RLCRQDGPAPVQERGAARMAVVAVFAISFLPFHTTKTAYLAV 265  
QY 255 RLDSWPQGTQKRAIKSIYTLTRPLAFNSAINPI-FYFLMGDHYREMLISKFRQ 307  
Db 266 RSTPGVCTVLEAFAAAYKTRPFASANSVLPILFYFTQ-----KKERR 310  
RESULT 13  
A45177  
chemokine (C-C) receptor 1 - human  
N:Alternate names: C-C CKR-1; macrophage inflammatory protein-1-alpha receptor  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 13-Aug-1999  
C:Accession: A45177; I55671  
R:Neote, K.; DiGregorio, D.; Mak, J.Y.; Horuk, R.; Schall, T.J.  
Cell 72, 415-425, 1993  
A:Title: Molecular cloning, functional expression, and signalling characteristics of a C-  
A:Reference number: A45177; MUID:93161416; PMID:7679328  
A:Accession: A45177  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-355 <NEO>  
A:Cross-references: GB:L10918; NID:9292416; PIDN:AAA36543.1; PID:9292417  
A:Experimental source: HL60 cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:124876)  
R:Gao, J.  
J. Exp. Med. 177, 1421-1427, 1993  
A:Title: Structure and functional expression of the human macrophage inflammatory 1 alpha  
A:Reference number: I55671; MUID:93240122; PMID:7683036  
A:Accession: I55671  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-355 <RES>  
A:Cross-references: GB:L10918; NID:9292416; PIDN:AAA36543.1; PID:9292417  
C:Genetics:  
A:Gene: GDB:CMKBR1; CMKR-1  
A:Cross-references: GDB:138446; OMIM:601159  
A:Map position: 3p21-3p21  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: disulfide bond; G protein-coupled receptor; glycoprotein; phosphoprotein; tr  
F:36-60/Domain: transmembrane #status predicted <TM1>  
F:71-91/Domain: transmembrane #status predicted <TM2>  
F:108-129/Domain: transmembrane #status predicted <TM3>  
F:147-171/Domain: transmembrane #status predicted <TM4>  
F:205-223/Domain: transmembrane #status predicted <TM5>  
F:240-264/Domain: transmembrane #status predicted <TM6>  
F:288-305/Domain: transmembrane #status predicted <TM7>  
F:5/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:24-273,106-183/Disulfide bonds: #status predicted  
F:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted  
Query Match 19.9%; Score 328; DB 2; Length 355;  
Best Local Similarity 29.0%; Pred. No. 6.1e-20;  
Matches 85; Conservative 63; Mismatches 127; Indels 18; Gaps 8;  
QY 22 LSAFYAIEFIFGLGNVTVPFGYLCFMKNWNSNVFLNLSISDFAFCLTLPILIKSYAN 81  
Db 37 LPPLYSLVFVIGLVGNILVLVLVQYRKLNKNTSIYLLNLAISDLLFLFTLPFWIDYK 96  
QY 82 DKGTGYDGLC--ISNRYVLHNLXTLSILFTLFIISMDRIYLMKYPFHEHLOKKEFAILIS 139  
Db 97 DDWFGDAMCKTLGSGFY--YTGLYSEIFFIILLTIDRYLAIVHAVALRARTVTFGVITS 154  
QY 140 LAVNALVLEVLPMLTFTNSVPKKEGSCNC-IDYASSGNPEHNLISLCLTLGLGLPLSV 198  
Db 155 IINWALIASMPGLYFFSKTQWEFTHHTSCSHFPESHREWKLFQALKNLFGVLVLLV 214  
QY 199 MCFEYKVMVFLKRRSQOQATAPLDPKQRLVLLAVVIFSLFTPYHIMRNLRASRLD- 257  
A:Residues: 1-388 <XUY>

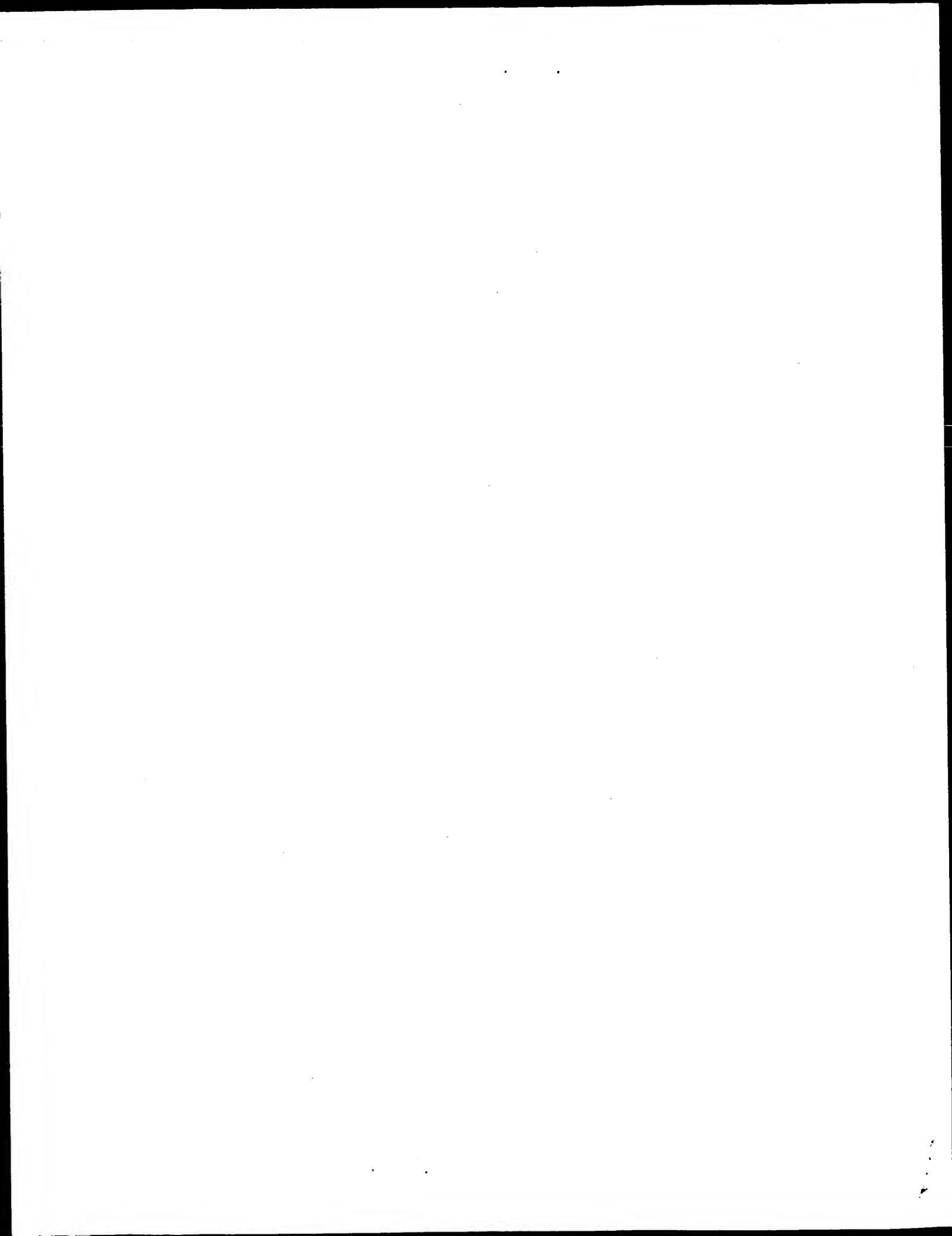
Db 215 MIICYGTGIKILLRRPNEKKS-----KAVRLFIVIMIIFFLFTWPNYL--TILISVFQDF 267  
QY 258 SWPOGCTQ-KAISKIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKPRQYF 309  
Db 268 LFTHECEQSRHLDAVQVTEVIATHTCCVNPVIYAFVGERFRKYL-----RQLF 316  
RESULT 14  
JC4587  
chemokine (C-C) receptor 4 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 08-Mar-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Jun-2000  
C:Accession: JC4587  
R:Hoogewerf, A.J.; Black, D.; Proudfoot, A.E.I.; Wells, T.N.C.; Power, C.A.  
Biochem. Biophys. Res. Commun. 218, 337-343, 1996  
A:Title: Molecular cloning of murine CC CKR-4 and high affinity binding of chemokines  
A:Reference number: JC4587; MUID:96136324; PMID:8573157  
A:Accession: JC4587  
A:Molecule type: mRNA  
A:Residues: 1-360 <HOO>  
A:Cross-references: EMBL:X90862; NID:g1167851; PIDN:CAA62372.1; PID:g1167852  
A:Experimental source: thymus  
C:Genetics:  
A:Gene: cc ckr-4  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: glycoprotein; phosphoprotein; receptor; thymus  
F:2,183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:72,202,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status p  
F:145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predi  
F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predi  
Query Match 19.8%; Score 326.5; DB 2; Length 360;  
Best Local Similarity 28.4%; Pred. No. 8.2e-20;  
Matches 93; Conservative 57; Mismatches 136; Indels 41; Gaps 10;  
QY 14 EAILNKYY-----LSAFYAIEFIFGLGNVTVPFGYLCFMKNWN 52  
Db 13 .EVVYNSYFYESNPKPCYKEGKAFGEVFLPPLYSLSVLLGLFGNSVVVL-VLFKYRKLK 71  
QY 53 S-SNYVFLNLSISDFAFCLTLPILIKSYANDKGTGYDGLVCISNRYVLHNLXTLSILFTTF 111  
Db 72 SMTDVLNLAISDLLFLVLSLPFW-GYYAADQWVEGLCKIVSNMYLVGYSGIFFIML 130  
QY 112 ISMDRYLLMKYFPEHFKQKKEFAILISLAVNALVTLEVLPMLTFTNSVPKKEGSCNCIDY 171  
Db 131 MSIDRYLAIVHAVALRARTVTFGVITSILTSVAVAFASLPOLLSTCYTEHNHTYCKTQ 190  
QY 172 ASSGNPEHNLISLCLTLGLFLPLSVCMCFYKVMVFLKRRSQOQATAPLDPKQRLV 231  
Db 191 YSVNSTTKVLSLEINVLGLLPLGIM-LFWYSMII-----RTLQHKNEKNRAVRMIF 245  
QY 232 LAVVIFSLFTPYHIMRNLRASRLDSNWPQGTCT-OKAISKIYTLTRPLAFNSAINPIFY 290  
Db 246 GVVLFGFTWPNVNVFLTLEVLVEVL-ODCTLERLYDIAQTATLGFTHCCLNPVIY 304  
QY 291 FLMGDHYREMLISKFRQYFKSLTSFT 317  
Db 305 FFLGE-----KFKYITQL--PRT 321  
RESULT 15  
JN0605  
somatostatin receptor 4 - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000  
C:Accession: JN0605; JN0762; A47457  
R:Xu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.  
Biochem. Biophys. Res. Commun. 193, 648-652, 1993  
A:Title: Molecular cloning and sequencing of a human somatostatin receptor, hSTR4.  
A:Reference number: JN0605; MUID:93290656; PMID:8512564  
A:Accession: JN0605  
A:Molecule type: DNA  
A:Residues: 1-388 <XUY>

A:Cross-references: GB:L14856; NID:g292499; PIDN:AAA36623.1; PID:g292500  
F:Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I  
Biochem. Biophys. Res. Commun. 195, 844-852, 1993  
A:Title: Cloning, functional expression and pharmacological characterization of a fourth  
A:Reference number: JN0762; MUID:93384611; PMID:8373420  
A:Accession: JN0762  
A:Molecule type: DNA  
A:Residues: 1-388 <YAM>  
A:Cross-references: GB:D16826; NID:g693907; PIDN:BAA04106.1; PID:g693908  
R:Rohrer, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, F.; Schule, R.  
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993  
A:Title: Cloning and characterization of a fourth human somatostatin receptor.  
A:Reference number: A47457; MUID:93248256; PMID:8483934  
A:Accession: A47457  
A:Molecule type: DNA  
A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>  
A:Cross-references: GB:L07833; NID:g307429; PIDN:AAA60565.1; PID:g307430  
A:Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBI:130858)  
C:Comment: This protein mediates the diverse actions of the tetradeptide somatostatin.  
C:Genetics:  
A:Gene: GDB:SSTR4  
A:Cross-references: GDB:202662; OMIM:182454  
A:Map position: 20p11.2-20p11.2  
A:Introns: #status absent  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; pho  
F:47-73/Domain: transmembrane #status predicted <TM1>  
F:84-109/Domain: transmembrane #status predicted <TM2>  
F:121-142/Domain: transmembrane #status predicted <TM3>  
F:162-184/Domain: transmembrane #status predicted <TM4>  
F:208-238/Domain: transmembrane #status predicted <TM5>  
F:257-284/Domain: transmembrane #status predicted <TM6>  
F:291-314/Domain: transmembrane #status predicted <TM7>  
F:24/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:119-198/Disulfide bonds: #status predicted  
F:161,253/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pr  
F:327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 19.7%; Score 325.5; DB 2; Length 388;  
Best Local Similarity 28.4%; Pred. No. 1.le-19;  
Matches 82; Conservative 66; Mismatches 110; Indels 31; Gaps 5;

QY 22 LSAFYAIEFIFGLGNVTYVFCYLCMKWNSSNVLYFNLSISDFALCTLPILIKSYAN 81  
DQ 50 IQCIYALVCLVGLGNALVIFILRYAKMKATNIYLLNLAVALDELFLSVFPVASSAAL 109  
QY 82 DRGTYGDVLCISNRXYLHTNLTSILFTFISMDRYLLMKYPFEHFLQKKEFALLISLA 141  
DB 110 RHWPFGSVLCRAVLSVDGLNMFYSVCLTVLSVDYVAVVHPRAATYRPSVAKLNLG 169  
QY 142 VVALVTLEVLPMFTINSVPKEGNCIDYASSGNPEHNLISYCLTLGLFLPLSVMCF 201  
DB 170 VMLASLLVTLPTAIFADTRPARGQAVACNLQWPHPAWSAVFVYVTFLLGLFLLPVLAIGL 229  
QY 202 FYKMW-----VFL-----KRSQQQATALPDKPQLVVLAVVVFISILTPHYHIMNL 250  
DB 230 CYLLIVGKRAVALRAGHQORREKKT-----RLVLVVVVVFLCWMFPYVQLL 281  
QY 251 R-IASRLDSWPGCTQKAIKSIYTLTRPLAFNLNSAINIFYFLMGDHYR 298  
DB 282 NLVWTSLDA-----TVNHVSLISYANSCANPILYGLSDNFR 319

Search completed: April 2, 2003, 15:18:04  
Job time : 26 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:32:31 ; Search time 16 seconds  
(without alignments)  
821.750 Million cell updates/sec

Title: US-09-891-138a-2

Perfect score: 1650

Sequence: 1 MAONLSCENWLATEAILNKY.....REMLISKPROVFKSLTSGRT 317

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	29.0	373	1	P2YR_RAT
2	475	28.8	373	1	P2YR_MOUSE
3	474	28.7	373	1	P2YR_BOVIN
4	474	28.7	373	1	P2YR_HUMAN
5	472.5	28.6	362	1	P2YR_CHICK
6	472.5	28.6	362	1	P2YR_MELGA
7	386.5	23.4	537	1	P2Y8_XENLA
8	379.5	23.0	328	1	P2Y3_CHICK
9	374	22.7	374	1	P2Y2_RAT
10	373.5	22.6	373	1	P2Y2_MOUSE
11	371.5	22.5	328	1	P2Y3_MELGA
12	363.5	22.0	377	1	P2Y2_HUMAN
13	354	21.5	361	1	P2Y4_RAT
14	353	21.4	355	1	C3X1_HUMAN
15	352	21.3	365	1	P2Y4_HUMAN
16	343	20.8	361	1	P2Y4_MOUSE
17	336.5	20.4	360	1	CKR4_HUMAN
18	336	20.4	328	1	P2Y6_RAT
19	333	20.2	308	1	P2Y5_CHICK
20	332.5	20.2	354	1	C3X1_RAT
21	332	20.1	328	1	P2Y6_HUMAN
22	329	19.9	355	1	CKR1_MACMU
23	328	19.8	355	1	CKR1_HUMAN
24	326.5	19.8	346	1	CLT2_HUMAN
25	326.5	19.8	360	1	CKR4_MOUSE
26	325.5	19.7	354	1	C3X1_MOUSE
27	325.5	19.7	388	1	SSR4_HUMAN
28	319.5	19.4	345	1	CLT2_PIG
29	319	19.3	353	1	APJ_XENLA
30	318.5	19.3	359	1	IL8B_MOUSE
31	317.5	19.2	355	1	GPRI_HUMAN
32	317	19.2	368	1	SSR2_BOVIN
33	316.5	19.2	384	1	SSR4_RAT

34	315.5	19.1	355	1	GPRI_MACMU	O97664 macaca mula
35	315.5	19.1	355	1	IL8A_RABIT	P21109 oryctolagus
36	315	19.1	355	1	CKR8_HUMAN	P51685 homo sapien
37	314	19.0	369	1	SSR2_HUMAN	P30874 homo sapien
38	314	19.0	369	1	SSR2_PIG	P34994 sus scrofa
39	314	19.0	370	1	P2Y9_HUMAN	O96677 homo sapien
40	313.5	19.0	353	1	IL8B_PANTR	Q28807 pan troglod
41	313.5	19.0	356	1	CKR8_MACMU	O97665 macaca mula
42	313.5	19.0	356	1	IL8B_CANFA	O97571 canis famil
43	313	19.0	344	1	P2Y5_HUMAN	P43657 homo sapien
44	312.5	18.9	360	1	IL8B_HUMAN	P25025 homo sapien
45	311.5	18.9	309	1	CLT2_MOUSE	Q920a1 mus musculu

ALIGNMENTS

RESULT 1						
P2YR_RAT						
ID	P2YR_RAT	STANDARD;	PRT;	373	AA.	
AC	P49651;					
DT	01-FEB-1996 (Rel. 33, Created)					
DT	01-FEB-1996 (Rel. 33, Last sequence update)					
DT	01-FEB-1996 (Rel. 33, Last annotation update)					
DE	P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).					
GN	P2RY1					
OS	Rattus norvegicus (Rat).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.					
OX	NCBI_TaxID=10116;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Insulinoma;					
RX	MEDLINE=95298025; PubMed=7779087;					
RA	Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;					
RT	"Cloning of rat and mouse P2Y purinoceptors."					
RL	Biochem. Biophys. Res. Commun. 211:211-218(1995).					
CC	-!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS					
CC	ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS					
CC	OF EXTRACELLULAR ATP ON INSULIN SECRETION.					
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.					
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN MUSCLE, HEART, LIVER, KIDNEY,					
CC	LUNG, BRAIN, SPLEEN, BUT NOT IN TESTIS.					
CC	-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>					
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).					
EMBL	U22830; AA91303.1;					
HSSP	P34996; IDDD.					
DR	InterPro; IPR000276; GPCR_Rhodpsn.					
DR	Pfam; PF00001; 7tm.1; 1.					
DR	PRINTS; PR00237; GPCRHHODOPSN					
DR	PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.					
DR	PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.					
KW	G-protein coupled receptor; Transmembrane; Glycoprotein.					
FT	DOMAIN 1 52					
FT	TRANSMEM 53 74					
FT	DOMAIN 75 87					
FT	TRANSMEM 88 109					
FT	DOMAIN 110 126					
FT	TRANSMEM 127 147					
FT	DOMAIN 148 166					
FT	TRANSMEM 167 188					
FT	DOMAIN 189 218					
FT	TRANSMEM 219 238					
FT	DOMAIN 239 265					
FT	TRANSMEM 266 285					
FT	TRANSMEM 285					



FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 304 328 7 (POTENTIAL).  
 FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 124 202 BY SIMILARITY.  
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 373 AA; 42321 MW; 60DF676287B5E648 CRC64;

Query Match 29.0%; Score 479; DB 1; Length 373;  
 Best Local Similarity 37.3%; Pred. No. 5.4e-25;  
 Matches 107; Conservative 57; Mismatches 115; Indels 8; Gaps 6;

QY 20 YLSAFYAIEFTGLGNVTVVFGYLCFMKNWSSNVFLNLSISDFAFCTLPILIKSY 79  
 DB 52 YLPAVYILVFTIGLGNVAIMWFVFMKPSGVSVMENALADFLYVLTLPALIFY 111  
 QY 80 ANDKG-TYGDVLCISNRYVLTNLTSLFTFISMDRYLLMKYPRFHFLOKKEFAILI 138  
 DB 112 FNKTDWIFGDVCKLQRFIFHVNLYGSLFTCISAHRYSGVYPLKSLGRKKNAIYV 171  
 QY 139 SLAVNALVTLVPLMLTFINS-VPEEGSNCIDYASSGNPEHNLIYSLCLTLGLFLPLS 197  
 DB 172 SVLVMLIVVAISPILFYSGTGIRKNKTVTCYDSTDEYLSYFYISMCTTVAMFCIPLV 231  
 QY 198 VMCFYKVMVFLKRSQQOATAPL-DKQRLVAVLVAVVIFSLFTPYHIMNLRASRL 256  
 DB 232 LILGCVGLIVRALYKDLNS---PLRRKSIYVILVTVFAVSYPHFVMTNLRARL 288  
 QY 257 D-SWPQGCT-QKAISYITLTPRLAFNSAINPIFYFLMGDHYREML 301  
 DB 289 DFQTEPCMDENDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 2  
 P2YR\_MOUSE STANDARD; PRT; 373 AA.  
 AC P49650;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).  
 GN P2RY1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Insulinoma;  
 RA MEDLINE=95298025; PubMed=7779087;  
 RA Tokuyama Y, Hara M, Jones E.M.C., Fan Z., Bell G.I.;  
 RT "Cloning of rat and mouse P2Y purinoceptors";  
 RL Biochem. Biophys. Res. Commun. 211:211-218(1995).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=129/Sv;  
 RA Leon C.;  
 RT "Thromboresistance in P2Y1 receptor knockout mice";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS  
 CC ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS  
 CC OF EXTRACELLULAR ATP ON INSULIN SECRETION.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U22829; AAA91302.1; -;  
 DR EMBL: AJ245636; CAB57317.1; -;  
 DR HSSP: P34996; 1DDO.  
 DR MGD; MGI:105049; P2ry1.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCRHHODPSN.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE: PS00262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 53 74 1 (POTENTIAL).  
 FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 109 2 (POTENTIAL).  
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 127 147 3 (POTENTIAL).  
 FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 167 188 4 (POTENTIAL).  
 FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 219 238 5 (POTENTIAL).  
 FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 266 285 6 (POTENTIAL).  
 FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 304 328 7 (POTENTIAL).  
 FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 124 202 BY SIMILARITY.  
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 373 AA; 42212 MW; 944125E9F4560BB3 CRC64;

Query Match 28.8%; Score 475; DB 1; Length 373;  
 Best Local Similarity 37.3%; Pred. No. 1e-24;  
 Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 7;

QY 20 YLSAFYAIEFTGLGNVTVVFGYLCFMKNWSSNVFLNLSISDFAFCTLPILIKSY 79  
 DB 52 YLPAVYILVFTIGLGNVAIMWFVFMKPSGVSVMENALADFLYVLTLPALIFY 111  
 QY 80 ANDKG-TYGDVLCISNRYVLTNLTSLFTFISMDRYLLMKYPRFHFLOKKEFAILI 138  
 DB 112 FNKTDWIFGDVCKLQRFIFHVNLYGSLFTCISAHRYSGVYPLKSLGRKKNAIYV 171  
 QY 139 SLAVNALVTLVPLMLTFINS-VPEEGSNCIDYASSGNPEHNLIYSLCLTLGLFLPLS 197  
 DB 172 SVLVMLIVVAISPILFYSGTGIRKNKTVTCYDSTDEYLSYFYISMCTTVAMFCIPL- 230  
 QY 198 VMCFYKVMVFLKRSQQOATAPL-DKQRLVAVLVAVVIFSLFTPYHIMNLRASRL 256  
 DB 231 VILGCVGLIV--KALYINDLNSPLRRKSIYVILVTVFAVSYPHFVMTNLRARL 288  
 QY 257 D-SWPQGCT-QKAISYITLTPRLAFNSAINPIFYFLMGDHYREML 301  
 DB 289 DFQTEPCMDENDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 3  
 P2YR\_BOVIN STANDARD; PRT; 373 AA.  
 AC P48042;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).  
 GN P2RY1.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;





SQ SEQUENCE 362 AA; 41194 MW; A806C88FB9514761 CRC64;

Query Match 28.6%; Score 472.5; DB 1; Length 362;  
 Best Local Similarity 35.9%; Pred. No. 1.4e-24;  
 Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;

QY 10 WLA-----TEALNK-----YLSAFYAIEFIFGLGNVTVVFGYLCMKMKNSSNVYLFN 60  
 DB 22 WAAGNATTKSLKTKGFQFYLLVITFTGLGNSVAIMFVPHMRPWSGISVYMFN 81

QY 61 LSSDFAFCLTLPILIKSYANDKG-TYGDVLCISNRVYLTHTNLYTSILFTFISMDRYLL 119  
 DB 82 LALADELYVLTLPALIFYFNKTDWIFGDVCMCKLQRFHFVNLYGSLFTLCISVHYRTG 141

QY 120 MKYFFREHFLOKKEFAILISLAWALVTLVLPMLTFINS-VPKEGSGNCIDYASSGNPE 178  
 DB 142 VVHPLKSLGLRKKNAVYVSSLVWALVAVIAPILFYSGTVGRNKITCYDTTADAYLR 201

QY 179 HNLTPYHIMRNLIARSLD-SWPGQCT-QKAISKIYITLTPPLAFNLAINPIFYFLMGD 295  
 DB 259 ANSVLPFHVMKTLNLRALDFQTPQMCAPNDKVYATYQVTRGLASLNSCVDPILYFLAGD 318

QY 296 HYREML 301  
 DB 319 TFRRL 324

RESULT 6  
 P2YR\_MELGA  
 ID P2YR\_MELGA STANDARD; PRT: 362 AA.  
 AC P49652;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (6H1 orphan receptor).  
 DE P2RY1.  
 OS Euleagris gallopavo (Common turkey).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.  
 OX NCBI\_TaxID=9103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=94335907; PubMed=8058061;  
 RA Filtz T.M., Li Q., Boyer J.L., Nicholas R.A., Harden T.K.;  
 RT "Expression of a cloned p2y purinergic receptor that couples to phospholipase C.";  
 RL Mol. Pharmacol. 46:8-14(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97382456; PubMed=9240460;  
 RA Li Q., Schachter J.B., Harden T.K., Nicholas R.A.;  
 RT "The 6H1 orphan receptor, claimed to be the p2y5 receptor, does not mediate nucleotide-promoted second messenger responses.";  
 RL Biochem. Biophys. Res. Commun. 236:455-460(1997).  
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: MAINLY FOUND IN BLOOD, BRAIN, AND LUNG. TO A LESSER EXTENT IN STOMACH, GUT, AND SKELETAL MUSCLE.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

-----  
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; U09842; AAA18784.1; -.  
 DR HSSP; P34996; 1DDD.  
 DR InterPro; IPR000276; GPCR\_Rhodospn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPS.  
 DR PROSITE; PS00237; G-PROTEIN RECP\_F1.1; 1.  
 DR PROSITE; PS0262; G-PROTEIN RECP\_F1.2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 41  
 FT TRANSFEM 42 63  
 FT DOMAIN 64 76  
 FT TRANSFEM 77 98  
 FT DOMAIN 99 115  
 FT TRANSFEM 116 136  
 FT DOMAIN 137 155  
 FT TRANSFEM 156 177  
 FT DOMAIN 178 207  
 FT TRANSFEM 208 227  
 FT DOMAIN 228 254  
 FT TRANSFEM 255 274  
 FT DOMAIN 275 292  
 FT TRANSFEM 293 317  
 FT DOMAIN 318 362  
 FT DISULFID 113 191  
 FT CARBOHYD 11 11  
 FT CARBOHYD 26 26  
 FT CARBOHYD 102 102  
 FT CARBOHYD 186 186  
 SQ SEQUENCE 362 AA; 41180 MW; 3E128AB9EB64349C CRC64;

Query Match 28.6%; Score 472.5; DB 1; Length 362;  
 Best Local Similarity 35.9%; Pred. No. 1.4e-24;  
 Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;

QY 10 WLA-----TEALNK-----YLSAFYAIEFIFGLGNVTVVFGYLCMKMKNSSNVYLFN 60  
 DB 22 WAAGNATTKSLTGTGFQFYLLVITFTGLGNSVAIMFVPHMRPWSGISVYMFN 81

QY 61 LSSDFAFCLTLPILIKSYANDKG-TYGDVLCISNRVYLTHTNLYTSILFTFISMDRYLL 119  
 DB 82 LALADELYVLTLPALIFYFNKTDWIFGDVCMCKLQRFHFVNLYGSLFTLCISVHYRTG 141

QY 120 MKYFFREHFLOKKEFAILISLAWALVTLVLPMLTFINS-VPKEGSGNCIDYASSGNPE 178  
 DB 142 VVHPLKSLGLRKKNAVYVSSLVWALVAVIAPILFYSGTVGRNKITCYDTTADAYLR 201

QY 179 HNLTPYHIMRNLIARSLD-SWPGQCT-QKAISKIYITLTPPLAFNLAINPIFYFLMGD 295  
 DB 259 ANSVLPFHVMKTLNLRALDFQTPQMCAPNDKVYATYQVTRGLASLNSCVDPILYFLAGD 318

QY 296 HYREML 301  
 DB 319 TFRRL 324

RESULT 7  
 P2Y8\_XENLA  
 ID P2Y8\_XENLA STANDARD; PRT: 537 AA.  
 AC P79928;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE P2Y purinoceptor 8 (P2Y8).  
 GN P2RY8.

OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Neural plate;  
 RX MEDLINE=97284734; PubMed=9139711;  
 RA Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;  
 RT "Early expression of a novel nucleotide receptor in the neural plate  
 of Xenopus embryos.";  
 RL J. Biol. Chem. 272:12593-12590(1997).  
 CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, CTP, GTP AND ITP.  
 CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH  
 CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.  
 CC MAY PLAY A KEY ROLE IN THE EARLY DEVELOPMENT OF NEURAL TISSUE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X99953; CAA68213.1; -;  
 DR HSP; P34996; 1DDD.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEPTOR\_FL\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECEPTOR\_FL\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 49  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 50 70  
 FT DOMAIN 71 79  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 80 100  
 FT DOMAIN 101 118  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 119 139  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 140 161  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 162 182  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 183 210  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 211 231  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 232 254  
 FT TRANSMEM 255 275  
 FT DOMAIN 276 292  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 293 316  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 317 337  
 FT BY SIMILARITY.  
 FT DISULFID 116 193  
 FT CARBOHYD 26 26  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 29 29  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SEQUENCE 537 AA; 62024 MW; B2CF24812F3C19F2 CRC64;  
 Query Match 23.4%; Score 386.5; DB 1; Length 537;  
 Best Local Similarity 35.3%; Pred. No. 9.2e-19;  
 Matches 104; Conservative 51; Mismatches 125; Indels 15; Gaps 7;  
 QY 22 LSAPYATEFTFGLGNVTVVGVFLCMKNWSSNVFLNLSISDFALCTLPILIKSYAN 81  
 DB 46 LPVSIVAFVWGLPNTAAMWIFIAKMRPNWPTVTFNLSLDTLVLSLPTLVYYA- 104  
 QY 82 DKGT--XGVDLCISNRVLTHTNLTSTLFTFISMDRYLMKYPFFREHFLQKEFAILIS 139  
 DB 105 DKNWNPGEVLCKLVRLFLFANLYSSILFTLCISVHRYGVCHPITSLRRMNAKHAYVIC 164  
 QY 140 LAVWALPTELVPLMLTINSVPKESGNCIDYASSGNPEHNLISCLTLGLPLPSVM 199  
 DB 165 ALVWLSVTLCVPLNLFVTPSKVKNITCHDTRPEDFARYVEYSTAIMCLLFGICPLII 224  
 QY 200 CFFYKXWVFLKRR--SQQATALPLDKPQ--RLVVLAVVIFSIETPTPYHIMNRLIASR 255

Db 225 AGCYGLMRELKMKPIVSGNQOT-LPSYKRSIKTIIFVMAFAICFMPHITRTLYYAR 283  
 QY 256 LDSWPOGCTQKAKSI---YTLTRPLAFNSAINPIFYFLMGDHYREMLISKFRQ 307  
 Db 284 L-----LGIKYALNVINVTYKVKTRPLASANSCLDPILYFLANDRYRRLIRVRR 334  
 [1]  
 RESULT 8  
 P2Y3\_CHICK  
 ID P2Y3\_CHICK STANDARD; PRT; 328 AA.  
 AC Q98907;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).  
 GN P2Y3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96319774; PubMed=8700132;  
 RA Bateson A.N., Burnstock G., Barnard E.A.;  
 RA Webb T.E., Henderson D., King B.F., Wang S., Simon J.,  
 RT "A novel G protein-coupled P2 purinoceptor (P2Y3) activated  
 preferentially by nucleoside diphosphates.";  
 RL Mol. Pharmacol. 50:258-265(1996).  
 CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADP > UTP > ATP = UDP. THE  
 CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE  
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X98283; CAA66930.1; -;  
 DR HSP; P34996; 1DDD.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEPTOR\_FL\_1; FALSE\_NEG.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECEPTOR\_FL\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 22  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 23 43  
 FT DOMAIN 44 57  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78  
 FT DOMAIN 79 96  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 97 117  
 FT DOMAIN 118 139  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 140 160  
 FT DOMAIN 161 189  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 190 210  
 FT DOMAIN 211 231  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 232 252  
 FT DOMAIN 253 275  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 276 298  
 FT DOMAIN 299 323  
 FT EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 5 5  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT DISULFID 94 172  
 FT BY SIMILARITY.  
 FT SEQUENCE 328 AA; 37586 MW; 7A3BF1C91F54FAAB CRC64;  
 Query Match 23.0%; Score 379.5; DB 1; Length 328;



OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93281707; PubMed=7685114;  
RA Lustig K.D., Shiao A.K., Brake A.J., Julius D.;  
RT "Expression cloning of an ATP receptor from mouse neuroblastoma  
cells";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:5113-5117(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96316177; PubMed=8739319;  
RA Enomoto K., Furiya K., Moore R.C., Yamagishi S., Oka T., Maeno T.;  
RT "Expression cloning and signal transduction pathway of P2U receptor  
in mammary tumor cells";  
RL Biol. Signals 5:9-21(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Head, and Liver;  
RC MEDLINE=21083660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Sakai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Breast tumor;  
RA Strausberg R.;  
RN Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP MUTAGENESIS OF LYS-107; ARG-110; HIS-262; ARG-265; LYS-289 AND  
ARG-292.  
RX MEDLINE=95181393; PubMed=7876172;  
RA Erb L., Garrad R.C., Wang Y., Quinn T., Turner J.T., Weisman G.A.;  
RT "Site-directed mutagenesis of P2U purinoceptors. Positively charged  
amino acids in transmembrane helices 6 and 7 affect agonist potency  
and specificity";  
RL J. Biol. Chem. 270:4185-4188(1995).  
CC -|- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that  
activate a phosphatidylinositol-calcium second messenger system.  
CC The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP  
CC = ADP.  
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -|- TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND  
CC BRAIN.  
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L14751; AAA39871.1; -;  
CC EMBL; S83099; AAB50735.1; -;  
CC EMBL; AK005013; BAB23746.1; -;

DR EMBL; AK017378; BAB30719.1; -;  
DR EMBL; BC006613; AAB06613.1; -;  
DR HSSP; P34996; 1DDD  
DR MGD; MGI:105107; P2ry2.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS00237; G-PROTEIN RECF\_1\_1; 1.  
DR PROSITE; PS02362; G-PROTEIN RECF\_2\_1; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 33 59 1 (POTENTIAL).  
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 71 93 2 (POTENTIAL).  
FT DOMAIN 94 110 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 111 129 3 (POTENTIAL).  
FT DOMAIN 130 152 4 (POTENTIAL).  
FT TRANSMEM 153 172 5 (POTENTIAL).  
FT DOMAIN 173 194 6 (POTENTIAL).  
FT TRANSMEM 195 220 7 (POTENTIAL).  
FT DOMAIN 221 246 8 (POTENTIAL).  
FT TRANSMEM 247 269 9 (POTENTIAL).  
FT DOMAIN 270 287 10 (POTENTIAL).  
FT TRANSMEM 288 309 11 (POTENTIAL).  
FT DOMAIN 310 373 12 (POTENTIAL).  
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT BY SIMILARITY.  
FT FT K->I: NO EFFECT ON RECEPTOR ACTIVATION.  
FT FT R->L: NO EFFECT ON RECEPTOR ACTIVATION.  
FT FT H->L: DECREASE IN RECEPTOR ACTIVATION.  
FT FT R->L: DECREASE IN RECEPTOR ACTIVATION.  
FT FT K->I: NO EFFECT ON RECEPTOR ACTIVATION.  
FT FT R->L: DECREASE IN RECEPTOR ACTIVATION.  
FT FT E -> D (IN REF. 2).  
FT FT S -> R (IN REF. 2).  
FT FT T -> N (IN REF. 2).  
FT FT V -> M (IN REF. 2).  
FT FT V -> L (IN REF. 3).  
FT FT D -> N (IN REF. 2).  
FT FT KDI -> PVV (IN REF. 2).  
SQ SEQUENCE 373 AA; 42174 MW; 590BBE502E41B3AE CRC64;  
  
Query Match 22.6%; Score 373.5; DB 1; Length 373;  
Best Local Similarity 33.5%; Pred. No. 4.7e-18;  
Matches 94; Conservative 50; Mismatches 130; Indels 7; Gaps 5;  
  
QY 19 KY-VLSAFYAIEFIFGLGNVTVVGYLFCMKNNSSNYLNLISDFAPLCTPLIK 77  
DB 32 KYVLLPVSYGVVCLNCLNVALYIFLCRLKTNASTYMFHLAVSDSLYASLPVY 91  
QY 78 SYA-NDKGTGVDVLCISNRYVHLNLTSLFTFISMDRVLLMKYPFREHFLQKKEAI 136  
DB 92 YYARGDHPFSTVLCKLVRFELFTNLYCSILFLTCISVHRCGLVRLPLHSLRWGRYAR 151  
QY 137 LISLAVWLVLEVLPMLETFNSVPKSGNSCIDYASGNDEHNLISLCLTLGLFLPL 196  
DB 152 RVAAGWVVLVLAQAPVLYFTTTSVRGTRTCHTSARELFSHFVAYSSVMGLLPAVPE 211  
QY 197 SVMCFYKVVVFKRRSQOQATALPDKPQ-RLVWLVVVFISLTPHYHIMNLRIAS 254  
DB 212 SVILCYVLMARELLKPAYGTGGPLPRAKRKSVRTIALVLAVFALCFELPFHVTTRYSE 271  
QY 255 RLDSWPOQC-TQAKTSYTYITRPLAFNLNSAINIFFLMG 294  
DB 273 R-SLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAG 310  
  
RESULT 11  
ID P2Y3\_MEIGA STANDARD; PRT; 328 AA.  
AC O93361;  
DT 16-OCT-2001 (Rel. 40, Created)



DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).  
 GN P2Y3.  
 OS Meleagris gallopavo (Common turkey).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.  
 OX NCBI\_TaxID=9103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98401046; PubMed=9730913;  
 RA Li Q., Olesky M., Palmer R.K., Harden T.K., Nicholas R.A.;  
 RT "Evidence that the p2y3 receptor is the avian homologue of the  
 mammalian p2y6 receptor.";  
 RL Mol. Pharmacol. 54:541-546(1998).  
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UDP> ADP = UTP. THE ACTIVITY  
 OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A  
 PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF069555; AAC23863.1; -;  
 DR HSP; P34996; 1DD0.  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; FALSE\_NEG.  
 DR PROSITE; PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 23 43 1 (POTENTIAL).  
 FT DOMAIN 44 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 2 (POTENTIAL).  
 FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 97 117 3 (POTENTIAL).  
 FT DOMAIN 118 139 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 140 160 4 (POTENTIAL).  
 FT DOMAIN 161 189 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 190 210 5 (POTENTIAL).  
 FT DOMAIN 211 231 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 232 252 6 (POTENTIAL).  
 FT DOMAIN 253 275 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 276 298 7 (POTENTIAL).  
 FT DOMAIN 299 323 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 5 5 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT DISULFID 94 172 BY SIMILARITY.  
 SQ SEQUENCE 328 AA; 37594 MW; B74D49B99C7164A5 CRC64;  
 Query Match 22.5%; Score 371.5; DB 1; Length 328;  
 Best Local Similarity 29.4%; Pred. No. 5.7e-18;  
 Matches 87; Conservative 61; Mismatches 137; Indels 11; Gaps 4;  
 QY 22 ISAFYAIFRGLGNVTWVFVGYLFCMKNSSNVVFLNLSISDFALCPLPLIKSYA- 80  
 24 LPLVTVSVFLLGLPLNAVVGIGWILQWALKALTRTIYMLNATADLLYVCSLPLIYNTQ 83  
 81 NDKGTGVDVLCISNRYLVHNNLYTSILFLTFISMDRYLLMKYPFRE-HFLQKKEAFILIS 139  
 84 KDYPFGDFCTKFRFQFYFTNLHGSLFLCISVQRYMGICHELASWHRKKGKLTWLC 143  
 140 LAVVALVTLVPLMLTFTINSVPKEGSGNCIDYASSGNPHNLIYSLCLLLGLFLPLSV 199  
 144 AAVFVIAOCLPTFFVFASTGTQNRNVTCVYDLSPPORSASYFPYGTLLITGFLFPFAI 203  
 200 CFFYKVMVFLKRRSQOQATAL--PLDKPQRLVVLAVVIFSLTPTPHIMRNLRITASRD 257

Db 204 LACYCSMARILCOKDELIGLAVHKKKAVRMIIIVIVFSISFPFHLTKIYLIIVSS 263  
 QY 258 SWPGCTQKAISLYITLRLPLAFNLNSAINPIFYFLMGDHYRE-----MLISKFR 306  
 Db 264 PTLPCPTLQAFATAYKCTRFPSMNSVLDPLFYFTQRFKRESTRYLDDKMSKWR 319  
 RESULT 12  
 P2Y2\_HUMAN  
 ID P2Y2\_HUMAN STANDARD; PRT; 377 AA.  
 AC P41231; Q96DM8;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)  
 DE (Purinoceptor 2) (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)  
 GN P2RY2 OR P2RU1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Airway epithelium;  
 RX MEDLINE=94211846; PubMed=8159738;  
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,  
 RA Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;  
 RT "Cloning and expression of a human P2U nucleotide receptor, a target  
 for cystic fibrosis pharmacotherapy.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 91:3275-3279(1994).  
 RN [2]  
 RP REVISIONS.  
 RX MEDLINE=95108098; PubMed=7809171;  
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,  
 RA Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;  
 RT "Cloning and expression of a human P2U nucleotide receptor, a target  
 for cystic fibrosis pharmacotherapy.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 91:13067-13067(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney, and Leukocyte;  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that  
 activate a phosphatidylinositol-calcium second messenger system.  
 CC The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP  
 = ADP.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: SPLEEN, KIDNEY, LIVER, LUNG, HEART AND  
 BRAIN.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U07225; AAC04923.1; -;  
 DR EMBL; BC012104; AAH12104.1; -;  
 DR EMBL; BC028135; AAH28135.1; -;  
 DR PIR; A54946; A54946.  
 DR HSP; P34996; 1DD0.  
 DR Genew; HGNC:8541; P2RY2.  
 DR MIM; 600041; -;  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.



[illegible]



SEQUENCE FROM N.A.  
MEDLINE=96125055; PubMed=8537336;  
Communi D., Piroton S., Parmentier M., Boeynaems J.-M.;  
"Cloning and functional expression of a human uridine nucleotide  
receptor.";  
J. Biol. Chem. 270:30849-30852(1995).  
[2]  
SEQUENCE FROM N.A.  
MEDLINE=96125054; PubMed=8537335;  
Nguyen T., Erb L., Welsman G.A., Marchese A., Heng H.H.Q.,  
Garrod R.C., George S.R., Turner J.T., O'Dowd B.F.;  
"Cloning, expression, and chromosomal localization of the human  
uridine nucleotide receptor gene.";  
J. Biol. Chem. 270:30845-30848(1995).  
[3]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas:  
MEDLINE=9617801; PubMed=8617367;  
Stam N.J., Klomp J., van der Heuvel M., Olijve W.;  
"Molecular cloning and characterization of a novel orphan receptor  
(P2p) expressed in human pancreas that shows high structural homology  
to the P2U purinoceptor.";  
FEBS Lett. 384:260-264(1996).  
CC -!- FUNCTION: Receptor for UTP and UDP coupled to G-proteins that  
activate a phosphatidylinositol-calcium second messenger system.  
CC - Not activated by ATP or ADP.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: Pancreas.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL: X91852; CAA62963.1; -;  
DR EMBL: U40223; AAC50347.1; -;  
DR EMBL: X96597; CRA65415.1; -;  
DR HSSP: F34996; IIDD.  
DR Genew; HGNC:8542; P2RY4.  
DR MIM; 300038; -;  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCRRHODOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECP\_FL\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECP\_FL\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Polymorphism.  
DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 35 61 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 62 72 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 73 95 2 (POTENTIAL).  
FT DOMAIN 96 112 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 113 131 3 (POTENTIAL).  
FT DOMAIN 132 154 4 (POTENTIAL).  
FT TRANSMEM 155 174 4 (POTENTIAL).  
FT DOMAIN 175 196 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 197 222 5 (POTENTIAL).  
FT DOMAIN 223 246 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 247 269 6 (POTENTIAL).  
FT DOMAIN 270 287 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 288 309 7 (POTENTIAL).  
FT DOMAIN 310 365 CYTOPLASMIC (POTENTIAL).  
FT FT DISULFID 108 185 BY SIMILARITY.  
FT VARIANT 168 168 /FTId=VAR\_011854.  
FT FT VARIANT 178 178 N -> T (IN DBSNP:1152187).  
FT FT VARIANT 191 191 P -> L (IN DBSNP:1152188).  
FT FT CONFLICT 86 86 /FTId=VAR\_011856.  
FT FT

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 15:16:42 ; Search time 42 Seconds  
(without alignments)  
1555.165 Million cell updates/sec

Title: US-09-891-138a-2  
Perfect score: 1650  
Sequence: 1 MAONLSCENWLATEAILNKY.....REMLISKFRQYFKSLTSFRT 317

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertibrate.\*
- 14: sp.unclassified.\*
- 15: sp.virus.\*
- 16: sp.bacteriap.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1634	99.0	317	11 Q99MT6	Q99mt6 mus musculus
2	1238.5	75.1	330	4 Q9BXA5	Q9bxa5 homo sapien
3	1238.5	75.1	334	4 Q8TDQ8	Q8tdq8 homo sapien
4	482.5	29.2	361	13 Q90X57	Q90x57 xenopus lae
5	461	27.9	357	13 Q9DE05	Q9de05 raja erinac
6	426	25.8	337	4 Q96P68	Q96p68 homo sapien
7	422.5	25.6	374	13 Q57466	Q57466 meleagris g
8	353	21.4	361	11 Q8VHP3	Q8vhp3 cavia porce
9	338	20.5	328	11 Q9ERK9	Q9erk9 mus musculus
10	335	20.3	355	6 Q9MTJ8	Q9myj8 callithrix
11	333	20.2	355	11 Q9JLY8	Q9jly8 rattus norv
12	326	19.8	351	6 Q9MTJ9	Q9myj9 oryctolagus
13	320	19.4	360	11 Q9IZH4	Q9izh4 rattus norv
14	318.5	19.3	309	11 Q8RS28	Q8rs28 mus musculus
15	316.5	19.2	355	6 Q95LH1	Q95lh1 macaca fasc
16	316	19.2	355	11 Q91VP9	Q91vp9 mus musculus

17	316	19.2	367	13 Q9PVG0	Q9pvq0 carassius a
18	315	19.1	355	4 Q9BYX5	Q9byx5 homo sapien
19	314	19.0	356	4 Q96TF2	Q96tf2 homo sapien
20	313	19.0	367	13 Q9PVF9	Q9pvf9 carassius a
21	312.5	18.9	346	4 Q96GE0	Q96ge0 homo sapien
22	312.5	18.9	390	13 Q8QGO4	Q8qgq4 carassius a
23	308	18.7	346	11 Q91W73	Q91w73 mus musculus
24	306.5	18.6	392	6 Q9BDQ4	Q9bdq4 canis famil
25	306	18.5	337	11 Q921N3	Q92ln3 mus musculus
26	301	18.2	380	13 Q9DGO6	Q9dgg6 carassius a
27	300	18.2	354	12 Q91OT6	Q91qt6 human cytom
28	299	18.1	346	6 Q95KS6	Q95ks6 ovis aries
29	299	18.1	354	12 Q91P65	Q91p65 human cytom
30	298	18.1	354	12 Q91Q00	Q91qu0 human cytom
31	298	18.1	354	12 Q91OK7	Q91ok7 human cytom
32	298	18.1	363	4 Q8TDS4	Q8tds4 homo sapien
33	297.5	18.0	358	6 Q9NOM0	Q9nm0 ovis aries
34	296	17.9	346	4 Q9BXC0	Q9bxc0 homo sapien
35	296	17.9	354	12 Q91P67	Q91p67 human cytom
36	296	17.9	354	12 Q91P66	Q91p66 human cytom
37	295	17.9	354	12 Q91P68	Q91p68 human cytom
38	295	17.9	354	12 Q91P69	Q91p69 human cytom
39	295	17.9	354	12 Q91QT9	Q91qt9 human cytom
40	295	17.9	354	12 Q91QT8	Q91qt8 human cytom
41	295	17.9	354	12 Q91QT7	Q91qt7 human cytom
42	295	17.9	354	12 Q91OH3	Q91oh3 human cytom
43	295	17.9	354	12 Q91OE0	Q91oe0 human cytom
44	294.5	17.8	360	13 Q8QPR6	Q8qpr6 xenopus lae
45	294	17.8	354	12 Q91QU2	Q91qu2 human cytom

ALIGNMENTS

RESULT 1  
Q99MT6 ID Q99MT6 PRELIMINARY; PRT; 317 AA.  
AC Q99MT6;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE G-protein coupled receptor GPR91.  
GN GPR91.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL;  
RX MEDLINE=21172992; PubMed=11273702;  
RA Wittenberger T., Schaller H.C., Hellebrand S.;  
RT "An expressed sequence tag (est) data mining strategy succeeding in  
the discovery of new g-protein coupled receptors.";  
RL J. Mol. Biol. 307:799-813(2001).  
DR EMBL; AF295367; AAK01867.1;  
DR HSSP; P34996; 1DDO.  
DR MGD; MGI:1934135; Gpr91.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
SQ RECEPTOR.  
KW SEQUENCE. 317 AA; 36701 MW; 4E22F0608F928E36 CRC64;  
Query Match 99.0%; Score 1634; DB 11; Length 317;  
Best Local Similarity 99.1%; Pred. No. 5.8e-138;  
Matches 314; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MAONLSCENWLATEAILNKYLSAFYAIFEGLLGNVTVFYGLFCMKNNSSNVYLFN 60  
Db 1 MAONLSCENWLATEAILNKYLSAFYAIFEGLLGNVTVFYGLFCMKNNSSNVYLFN 60

Thu Apr 3 10:42:38 2003

```

QY 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
Db 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
QY 121 KYPREHFLQKKEFAILISLAWALVTLVPLMLTFINSVPKEGSCNIDYASSGNPEHN 180
Db 121 KYPREHFLQKKEFAILISLAWALVTLVPLMLTFINSVPKEGSCNIDYASSGNPEHN 180
QY 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATAPLDPKQRLVVLAVVIFSIL 240
Db 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATAPLDPKQRLVVLAVVIFSIL 240
QY 241 FPTYHIMRNLRITASLDSWPQCTQKAIKSIYTLRPLAFNLAINPFIYFLMGDHYRE 300
Db 241 FPTYHIMRNLRITASLDSWPQCTQKAIKSIYTLRPLAFNLAINPFIYFLMGDHYRE 300
QY 301 LISKPROYFKSLTSF 317
Db 301 LISKPROYFKSLTSF 317

RESULT 2
Q9BXA5 PRELIMINARY; PRT; 330 AA.
AC Q9BXA5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE G-protein coupled receptor 91.
GN GPR91.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
RT the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF348078; AAK29080.1; -.
DR HSP; P34996; 1DD0.
DR InterPro; IPR002276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1. 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 330 AA; 38283 MW; 6D3C70E340866BAB CRC64;

```

```

Query Match 75.1%; Score 1238.5; DB 4; Length 330;
Best Local Similarity 72.2%; Pred. No. 1.2e-102;
Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNYLFN 60
Db 1 MAWNATCKNMLAAEALEKYISFYIGFVGVGLNTIVVYGYFISLKNWSSNYLFN 60
QY 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
Db 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
QY 121 KYPREHFLQKKEFAILISLAWALVTLVPLMLTFINSVPKEGSCNIDYASSGNPEHN 180
Db 121 KYPREHFLQKKEFAILISLAWALVTLVPLMLTFINSVPKEGSCNIDYASSGNPEHN 180
QY 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATAPLDPKQRLVVLAVVIFSIL 240
Db 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATAPLDPKQRLVVLAVVIFSIL 240

```

```

QY 241 FPTYHIMRNLRITASLDSWPQ-GCTQKAIKSIYTLRPLAFNLAINPFIYFLMGDHYRE 299
Db 241 FPTYHIMRNLRITASLDSWPQ-GCTQKAIKSIYTLRPLAFNLAINPFIYFLMGDHYRE 299
QY 300 MLISKPROYFKSLTSF 315
Db 301 MLMNQLRHNFKSLTSF 316

RESULT 3
Q8TDQ8 PRELIMINARY; PRT; 334 AA.
AC Q8TDQ8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE P2Y purinoceptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
Zhang W., Li N., Wan T., Cao X.;
RA "Human P2Y purinoceptor 1.";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247785; AAL95690.1; -.
DR EMBL; AF247785; AAL95690.1; -.
SQ SEQUENCE 334 AA; 38697 MW; 33146E1AD87F0E81 CRC64;

Query Match 75.1%; Score 1238.5; DB 4; Length 334;
Best Local Similarity 72.2%; Pred. No. 1.2e-102;
Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNYLFN 60
Db 5 MAWNATCKNMLAAEALEKYISFYIGFVGVGLNTIVVYGYFISLKNWSSNYLFN 64
QY 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
Db 65 LSVSDLAFLCTLPMLIRSYANGNIYGDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 124
QY 121 KYPREHFLQKKEFAILISLAWALVTLVPLMLTFINSVPKEGSCNIDYASSGNPEHN 180
Db 125 KYPREHFLQKKEFAILISLAWALVTLVPLMLTFINSVPKEGSCNIDYASSGNPEHN 184
QY 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATAPLDPKQRLVVLAVVIFSIL 240
Db 185 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATAPLDPKQRLVVLAVVIFSIL 244
QY 241 FPTYHIMRNLRITASLDSWPQ-GCTQKAIKSIYTLRPLAFNLAINPFIYFLMGDHYRE 299
Db 245 FPTYHIMRNLRITASLDSWPQ-GCTQKAIKSIYTLRPLAFNLAINPFIYFLMGDHYRE 304
QY 300 MLISKPROYFKSLTSF 315
Db 305 MLMNQLRHNFKSLTSF 320

RESULT 4
Q90X57 PRELIMINARY; PRT; 361 AA.
AC Q90X57;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE P2Y1 nucleotide receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]

```

```

Best Local Similarity 35.0%; Pred. No. 3.3e-33;
Matches 110; Conservative 64; Mismatches 112; Indels 28; Gaps 11;

QY 17 LNK-----YYLSAFYAIEIFGLLGNTVVVFGYLCMKNNSSNVYLFNLISDFAPLCTL 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 LNKGFQFYPLIMYIIVFTVFGIGNSVALWFIPIHMRPSSITIMYFNLVLADLFYVFSL 88

QY 73 PILIKSYANDKG--TYGDVLCISNRYLVHTNLTYSILFTFISMDRYLLMKYPFREFHLOK 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 89 PILIFYFNKTDWIFEGELCKLXRFIFHVNLYGSLFTLCISVHRVTGVVHPMKSLGRK 148

QY 132 KEFALLISLWALVTLEVLPMFLFINS-VPEKGSNCIDVASSGNEHNLIIYSLCLTLL 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 KSAIIVCVCWITVMAGISPLIFESRTGLRRNTICYDTTSKELLTFIYISMSTFF 208

QY 191 GELIPLS--VMCF-FYKMMVFLKRRSQOQALPLDKPQRLVVLVAVYFISILFTPYHIM 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 GCIPCFATILCYGFIKVALI-----SNDKMTPL-RCKSVRLVIIVLAVFAISVLPFHV 262

QY 248 NLRITASRLDSWQPG--CT-QKAIKSIYLTITRPLAFLNSAINPIFFFLMGDHYREMLIS 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 KNLNLQSLR--YYQGLDTCENRRRYATYQVIRGLASLNSCVDPIYFLAGDTFR----- 315

QY 304 KFRQYKSLTSFT 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 --RRETNAASREMT 327

RESULT 6
Q96P68 PRELIMINARY; PRT; 337 AA.
AC Q96P68;
DT 01-DEC-2001 (TrEMBLrel. 19, Created);
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update);
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update);
DE G protein-coupled receptor (Putative G-protein coupled receptor).
GN GPR80 OR GPCR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21438557; PubMed=11574155;
RA Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RA "Discovery and mapping of ten novel G protein-coupled receptor
RT genes.";
RL Gene 275:83-91(2001).
[2]
RN SEQUENCE FROM N.A.
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RP "Identification of G protein-coupled receptor genes from the human
RT genome sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411109; AAL26480.1; -.
DR EMBL; AB083598; BAB89311.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 337 AA; 38251 MW; 6814EA0044756CE6 CRC64;

Query Match 25.8%; Score 426; DB 4; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.1e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYYLSAFYAIEIFGLLGNTVVVFGYLCMKNNSSNVYLFNLISDFAPLCTLPILI 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31 LKWHYLPVIYGIIFVFCGNNAVISTYIFKMRPKWSSITIMLNACFDLLYLSLPLI 90

QY 77 KSYAN-DKGTGYDVLICSNRYLVHTNLTYSILFTFISMDRYLLMKYPFREFHFKKEFA 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 91 HYYASGENWIFGDFMCKFIRESFHNLYSSILFTCTSFYFCVLIHPMSCFSIHKTRCA 150
Qy 136 ILISLAVNALVTLEVPMLTINSVPKEEGSNCDIYASSGNPEHNLI--YSLCLTLTLGLFL 193
Db 151 VVACAVWLIISIVAPITPMTLITSTNRTNSACLDLTSS--DELNTIKWYNLILTAATTC 208
Qy 194 IPLSVCMFFYKVMVFLKRRSQOQATPLDKPQRLVVLAVVIFSLFTPHIMRNLRJA 253
Db 209 LPLVITVLCYTTIIHTHGLQDSC--KQARRLTLILLALAFYVCFPLPHILIRIE 266
Qy 254 SRLDSWPGOCT-OKAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKFR 306
Db 267 SKLLS--ISCSIEHQIHEAYIVSRPLAALNTFGLNLLVYVSDNFQQAQVCTVR 318

RESULT 7
O57466 PRELIMINARY; PRT; 374 AA.
AC O57466;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE G protein coupled P2Y nucleotide receptor.
GN TP2Y.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
SEQUENCE FROM N.A.
RF*
RC TISSUE=BLOOD;
RX MEDLINE=98086419; PubMed=9415702;
RA Boyer J.L., Waldo G.L., Harden T.K.;
RT "Molecular cloning and expression of an avian G protein-coupled P2Y
receptor.";
RL Mol. Pharmacol. 52:928-934(1997).
DR EMBL; AF031897; AAC60339.1; -.
DR HSP; P34996; 1DDO.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PRO0237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 374 AA; 42594 MW; 849C465722BD02B CRC64;

Query Match 25.6%; Score 422.5; DB 13; Length 374;
Best Local Similarity 33.7%; Pred.No. 9.4e-30;
Matches 106; Conservative 59; Mismatches 119; Indels 31; Gaps 8;

Qy 10 WL-----APEA--ILNKYY----LSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNVY 57
Db 20 WLGNITAAEAKCVNEEFKILLPISYGVFVGLPLNSWAMWIFVSRMPWNATTTY 79

Qy 58 LFNLSISDFALCTLPILIKSYAN-DKGYGDVLCISNRVYVLTNLYTSILFTFETSMDR 116
Db 80 MFNLAISDTLYVFSLETLVYVYADRNWPFKGVKFCIVRFLFYANLYSSILFTCLSVHR 139

Qy 117 YLLMKYFPEHFLQKKEFALLISLAVNALVTLEVPMLTINSVPKEEGSNCDIYASSGN 176
Db 140 YMGICHPIRSKWKTKHARLCIGVNLVVTICLIENLIFVTSSKDNSTLCHDTTKPEE 199

Qy 177 PEHNLVSLCLTLGLFPLISVMCFYKVMVFLKRRSQOQATPLDKPQ-----227
Db 200 FDVHVHYSISSIMALLFGIPFLVIVWCYCLMAKRLCKRS-----FPSSPRPVPSYKRSI 253

Qy 228 RLVLAVVIFSLFTPHIMRNLRIRASRLDSWPGQC-TQKAISYITLTPPLAFNSAIN 286
Db 254 KMIIIVTFVFAICVPVPHITRTLYTSRY--FOADCOTLIINFTYKTRPLASINSCLD 311

Qy 287 PIFVFLMGDHYREML 301
Db 312 PILYFMAGDKYRGRL 326
```

```
RESULT 8
Q8VHP3 PRELIMINARY; PRT; 361 AA.
AC Q8VHP3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CC-chemokine receptor 4.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A.
RA Jopling L.A., Sabroe I., Andrew D.P., Mitchell T.J., Li Y.,
RA Hodge M.R., Williams T.J., Pease J.E.;
RT "The identification, characterization and distribution of guinea pig
CCR4 and epitope mapping of a blocking antibody.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF431971; AAL57488.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PRO0237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 361 AA; 41064 MW; 9304E897D4FD6839 CRC64;

Query Match 21.4%; Score 353; DB 11; Length 361;
Best Local Similarity 30.0%; Pred.No. 1.5e-23;
Matches 86; Conservative 62; Mismatches 131; Indels 8; Gaps 5;

Qy 21 YLSAFYAIEFIFGLGNVTVVFGYLCMKNNWS-SNVYFLNLSISDFALCTLPILIKSY 79
Db 42 FLPLYSILVFLGLGNSVVVL-VLFYKRLSRMTDVLNLAISDLFLVLSLPFW-GYY 99

Qy 80 ANDKGTGVDVLCISNRVYVLTNLYTSILFTFISMDRYLLMKYFPREHFLQKKEFALLIS 139
Db 100 AADQWVFLGVLGCKWISYIYLVGYSIGIFFVILMSIDRYLAIVHGVFSMRVTTYGVITS 159

Qy 140 LAVNALVTLEVPMLTINSVPKEEGSNCDIYASSGNPEHNLVYSLCLTLGLFPLISVM 199
Db 160 LATWAVAVFASLPLGLFSTCYTERTNHTSCKTRYGSANSTWKLSSLEINILGLVPLIGM 219

Qy 200 CFFYKVMVFLKRRSQOQATPLDKPQRLVVLAVVIFSLFTPHIMRNLRIRASRLDSW 259
Db 220 -LFCYSMTII---RTLQCKSKKKNKAVKMFVAVVVLFGFWTPYINVLFTYTLVELEVL 274

Qy 260 PQGCTOKAISKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKFR 306
Db 275 QDCSLEKYLDFALQNTETLAFIHCLNLPITYFFLGEKFRKIVQLFK 321

RESULT 9
Q9ERK9 PRELIMINARY; PRT; 328 AA.
AC Q9ERK9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE P2Y6 receptor (Hypothetical 36.7 kDa protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA STRAIN=129/SVEV;
RX MEDLINE=21160052; PubMed=11259526;
RX Lazarowski E.R., Rochelle L.G., O'Neal W.K., Ribeiro C.M.P.,
RX Grubb B.R., Zhang V., Harden T.K., Boucher R.C.;
```



RT "Cloning and functional characterization of two murine uridine  
 RT nucleotide receptors reveal a potential target for correcting ion  
 RL transport deficiency in cystic fibrosis gallbladder.";  
 RN J. Pharmacol. Exp. Ther. 297:43-49(2001).  
 RP SEQUENCE FROM N.A.

RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF298899; AAG24619.1; -;  
 DR EMBL; BC027331; AAH27331.1; -;  
 DR HSP; P34996; IDDP;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSN.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEPT\_F1\_2; 1.  
 KW Hypothetical protein; Receptor.  
 SQ SEQUENCE 328 AA; 36721 MW; 00F9DF5ADADF903E CRC64;

Query Match 20.5%; Score 338; DB 11; Length 328;  
 Best Local Similarity 29.6%; Pred. No. 2.9e-22;  
 Matches 84; Conservative 54; Mismatches 136; Indels 10; Gaps 4;

QY 22 LSFAFAIEFIFGLGNVTGVVFGYLCMKNNSSNVYLFNLSISDFALCTLPILIKSYA- 80  
 DB 29 LTPVSVLVVGLPNCIVIAICASRRTLTSTRVYVTLNLAADLMYACSLPLIYNVAR 88  
 QY 81 NDKGTGYDVLCSNRYVLTHTNLTYSILFTFISMDRYLMLKYPFRE-HFLQKKEFAILIS 139  
 DB 89 GDHWPFGDLACRFVRLFYANLHGSILFTCISFQRYGICHPLASWHKRGRRAAWVVC 148  
 QY 140 LAVWALVTLEVLPMLTFTINSVPKKEGNCIDYASSGNPEHNLIYSLCLTLGLFLPLSV 199  
 DB 149 GYVWLAVTAAQCLTFAVFAATGIGTQRNRTVCYDLSPPILSTRVLPYGMALTVIGFLPF 208  
 QY 200 CFYIKVNVFLKRRSQQAATLPL-----DKPQLVWLVAVVIFSLFTFYHIMNRLTAS 254  
 DB 209 LACYERMA---RLRCRQGPAGVQAQRSSKAARMAVVAVFALSLFPFHTITAYLAV 265  
 QY 255 RLDSWPOGCTOKAISKIYTLRPLAFLNSAINPIFYFLMGDHYR 298  
 DB 266 RSTPGVSCPVLTEFAAYKGTGTRPFASVNSVLDPLILFYFTQOKFR 309

RESULT 10  
 Q9MJY8 PRELIMINARY; PRT; 355 AA.  
 AC Q9MJY8;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE Chemokine receptor.  
 GN CCRI.  
 OS Callithrix jacchus (Common marmoset).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.  
 OX NCBI\_TaxID=9483;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20153429; PubMed=10686294;  
 RA Liang M., Rosser M.P., Ng H.P., May K., Bauman J.G., Islam I.,  
 RA Ghannam A., Kretschmer P.J., Pu H., Dunning L., Snider R.M.,  
 RA Morrissey M.M., Hesselgesser J., Perez H.D., Horuk R.;  
 RT "Species selectivity of a small molecule antagonist for the CCR1  
 RT chemokine receptor.";  
 RL Eur. J. Pharmacol. 389:41-49(2000).  
 DR EMBL; AF127528; AAF36453.1; -;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEPT\_F1\_1; UNKNOWN\_1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEPT\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 355 AA; 40928 MW; 2B01C47E9874A2C1 CRC64;

Query Match 20.3%; Score 335; DB 6; Length 355;  
 Best Local Similarity 28.1%; Pred. No. 5.8e-22;  
 Matches 85; Conservative 63; Mismatches 138; Indels 16; Gaps 7;

QY 12 ATEAILNKYLSAFYAIEFIFGLGNVTGVVFGYLCMKNNSSNVYLFNLSISDFALCT 71  
 DB 27 ANERAFGAKLLPPLYSLVFVIGLVGNILVVVVOYKRLKNMTSIYLNLAISDLLEFT 86  
 QY 72 LPILIKSYANDKGTGYDVLCSNRYVLTHTNLTYSILFTFISMDRYLMLKYPFREHFL 129  
 DB 87 LPFWISYQLKTDWVFGNAMKVLVSFY--YTGLYSEIFFIILLIDRYLAIVHAVPALRA 144  
 QY 130 QKKKEFAILISLAVWALVTLEVLPMLTFTINSVPKKEGNC-IDYASSGNPEHNLIYSLCLT 188  
 DB 145 RTVTFGVITISIIWVLAIALSLPGLYFAKTQWEITHRTCSLHFPHEGSRQEWKLFQALKN 204  
 QY 189 LLGFLIPLSVVMCFYKVMVFLKRRSQQAATLPLDKPQLVWLVAVVIFSLFTFYHIMR 248  
 DB 205 LLGLVLPVLVMTVCTGIKILLRPNKKS-----KAVRLFVIMIIFFLFTWTPYNTT 259  
 QY 249 NLRIASRLDSWPOGCTO-KAISKIYTLRPLAFLNSAINPIFYFLMGDHYREMLISKFRQ 307  
 DB 260 LISVFQDF-LFTYCEQGRQLDLAIQVETMAYTHCCVNPVIYAFVGERFKHL----RQ 314  
 QY 308 YF 309  
 DB 315 LF 316

RESULT 11  
 Q9JLY8 PRELIMINARY; PRT; 355 AA.  
 AC Q9JLY8;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Macrophage inflammatory protein-1 alpha receptor.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD;  
 RX MEDLINE=20555330; PubMed=11091494;  
 RA Waller A., Nayee P., Czaplowski L.G.;  
 RT "Identification and characterization of a rat macrophage inflammatory  
 RT protein-1 alpha receptor.";  
 RL J. Hematother. Stem Cell Res. 9:703-710(2000).  
 DR EMBL; AF119381; AAF34340.1; -;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEPT\_F1\_1; UNKNOWN\_1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEPT\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 355 AA; 40838 MW; 2FEB8661D1E6E075 CRC64;

Query Match 20.2%; Score 333; DB 11; Length 355;  
 Best Local Similarity 28.6%; Pred. No. 8.7e-22;  
 Matches 85; Conservative 69; Mismatches 119; Indels 24; Gaps 9;

QY 22 LSFAFAIEFIFGLGNVTGVVFGYLCMKNNSSNVYLFNLSISDFALCTLPILIKSYAN 81  
 DB 37 LPLYSFVFIIGVGNILVILMOHRRQSMTSIYLFNLAVSDLVFLFTFPWDYKIK 96  
 QY 82 DKGTGYDVLCSNRYVLTHTNLTYSILFTFISMDRYLMLKYPFREHFLQKKEFAILIS 139  
 DB 97 DNWVFGDAMKLLSGFYLL--GLYSEIFFIILLIDRYLAIVHAVFSLRPARVITGILTS 154  
 QY 140 LAVWALVTLEVLPMLTFTINSVPKKEGNCIDYASSGNPEHNLIYSLCLTLGLFLI 194



```

Db 155 IIAWALAILASIPALCFKQWETHHTC-----SPHPDESILTKWRQALKNLGLIL 210
QY 195 PLSVMCFYKMW-VFLKRRSQOQATALPLDKPQRLVAVVIFSILFTPYHIMNRRIA 253
Db 211 PLLVMICVAGIIRILLRPNRKAQV-----RLFAITLFLLELLTWPYNL--TVFVS 262
QY 254 SRLD-SWFGQCTQ-KAISKIYTLTPFLAFLNSAINPIFYFLMGDHYREMLISKPROX 308
Db 263 AFQDVLFTNQCEQSKOLDLAIQVTEVIAIYTHCCVNPVIYVGERFKYLRQLFQRH 319

RESULT 12
Q9MYJ9 PRELIMINARY; PRT; 351 AA.
AC Q9MYJ9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Chemokine receptor.
GN CCRL.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20153429; PubMed=10686294;
RA Liang M., Rosser M.P., Ng H.P., May K., Bauman J.G., Islam I.,
RA Ghannam A., Kretschmer P.J., Pu H., Dunning L., Snider R.M.,
RA Morrissey M.M., Hesselgesser J., Perez H.D., Horuk R.;
RT "Species selectivity of a small molecule antagonist for the CCRL
RT chemokine receptor."
RL Eur. J. Pharmacol. 389:41-49(2000).
DR EMBL; AF127527; AAF36452.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1.1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 351 AA; 40534 MW; B3FED6117A141552 CRC64;

Query Match 19.8%; Score 326; DB 6; Length 351;
Best Local Similarity 27.8%; Pred. No. 3.6e-21;
Matches 83; Conservative 65; Mismatches 137; Indels 14; Gaps 7;

QY 22 LSFAFYAIEIFGLGNVTGVVFGVLCMKNNWSNVYLFNLISDFAFCTLPILIKSYAN 81
Db 37 LPPLYSLVFVIGLVGNVVLVLMKYKRLRSMTSIYLLNLAISDLLFLFPLFWIDYRLK 96
QY 82 DKGTGVDVLC--ISNRYVLHTNLTYSILFTFISMDRYLLMKYPFRHFLOKKEFALLIS 139
Db 97 DDWFGDVLCKFSLGLY--YVGLYSEVFVILLITIDRYLAIVHAVFALRARTVSGIVTS 154
QY 140 LAVWALVTLEVLPMLTFTINSVPKEGNSC-IDYASSGNPEHNLISYLCITLGLFLPLSV 198
Db 155 IVTWALTILAAIPGRFSKQWTFHTYCTSLHPHESLRQWQFOQALKNLILGLPLIV 214
QY 199 MCFEYKMWVFLKRRSQOQATALPLDKPQRLVAVVIFSILFTPYHIMNRRIASRLDS 258
Db 215 MVVYCTGTGIIQLLRNPKS-----RAVLFIIVMLIFLFWTPYNL--TLVSAFQDS 267
QY 259 -WPGQCTQ-KAISKIYTLTPFLAFLNSAINPIFYFLMGDHYREMLISKPROVFKSLTSF 315
Db 268 LFTNQCEQSKOLDLAIQVTEVIAIYTHCCVNPVIYVGERFQYLRQLFHTYLAQWLPF 326

RESULT 13
Q91ZH4 PRELIMINARY; PRT; 360 AA.
ID Q91ZH4;
AC Q91ZH4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

```

```

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE C-C chemokine receptor 4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW;
RA Garcia G.E., Chen S., Xia Y., Harrison J., Wilson C.B., Johnson R.J.,
RA Bacon K.B., Feng L.;
RT "Mononuclear cell-infiltrate inhibition by blocking macrophage-derived
RT chemokine results in attenuation of developing crescentic
RT glomerulonephritis.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432872; AAL30398.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1.1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 360 AA; 41218 MW; 5095C6CD299EIF8B CRC64;

Query Match 19.4%; Score 320; DB 11; Length 360;
Best Local Similarity 29.4%; Pred. No. 1.3e-20;
Matches 88; Conservative 58; Mismatches 133; Indels 20; Gaps 9;

QY 21 YLSAFYAIEIFGLGNVTGVVFGVLCMKNNWS-SNVYLFNLISDFAFCTLPILIKSY 79
Db 41 FLPLNSELVLLGLFGNSVVL-VLFYKRLKSWTDVYLLNLALISDLLFLVSLPFW-GY 98
QY 80 ANDKGTGVDVLCISNRYVLHTNLTYSILFTFISMDRYLLMKYPFRHFLOKKEFALLIS 139
Db 99 AADQWVEGLGLCKIISWMLVGVFGYGIFFILMSIDRYLAIVHAVFSLRARTLYGVITS 158
QY 140 LAVWALVTLEVLPMLTFTINSVPKEGNSCIDYASSGNPEHNLISYLCITLGLFLPLSV 199
Db 159 LITWSAVAFASLPGLLFTCDTENHHYCKYQTSVNSTTWKVLSSLEINVLGLVPLGIM 218
QY 200 CFYFKMWVFLKRRSQOQATALPLDKPQRLVAVVIFSILFTPYHIMNRRIASRLDSW 259
Db 219 -LFCYSMLIETLRHKNEKK-----NRAVRMTFAVVVFLGFWTPYNNVFLLEVL 273
QY 260 PQGCT-QKAISKIYTLTPFLAFLNSAINPIFYFLMGDHYREMLISKPROVFKSLTSPT 317
Db 274 -QDCTERYLDYAIQATETATIAFHCCLNPNVIYFLGE-----KFKYIAQL--PRT 321

RESULT 14
Q8R528 PRELIMINARY; PRT; 309 AA.
ID Q8R528;
AC Q8R528;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cysteineyl leukotriene 2 receptor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishii S.;
RT "Mouse CysLT2 Gene.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB058930; BAB86881.1; -.
KW Receptor.
SQ SEQUENCE 309 AA; 35226 MW; 96FACC6B8AF96974 CRC64;

Query Match 19.3%; Score 318.5; DB 11; Length 309;
Best Local Similarity 27.3%; Pred. No. 1.5e-20;
Matches 84; Conservative 56; Mismatches 147; Indels 21; Gaps 8;

```

```
QY 3 QNLSCENWLEATEILNKYLSAFVAIEFIFGLLGNTVVFVGLFCMKNNSSNVYLFNLS 62
Db 13 RNCIENF-----KKEFPYIYLIIFWALGNGFSYVFLQCKKSTSVNFMNLA 65
QY 63 ISDFAFICTLPILIKSY-ANDKGTGVDVLCISNRYVLTHTNLTSLFTFISMDRYLLMK 121
Db 66 TSDFLFISTLPFRADYFGRSNNWIFGLACRVMSYSLYVNMVTSYFLVAVLSVVRFLATV 125
QY 122 YPREHEFLQKKEFAILISLAVWALVTLVPLMLTFINSVPKKEGNS---CIDYASSGNPE 178
Db 126 HPRFMFHTSVRSAILWGLGIWIFMAS--SALLVNG--QEEKDNIIISCLE--LSPQKFK 180
QY 179 HNLISLCLTLGLFLPLSVKCFYFYMV-VFLKRRSQOQATALPLDKPQRLWVLAVVTF 237
Db 181 SLLIMNHIAVAGFLLPFLTLTICYLIIIRILLKAIPESGPRAHRKALTIVIAMITF 240
QY 238 SILFTPYHIMRNRIASRLDSWQDQCTQKAISYITLTPRLAFLNSAINPIFYFLMGDHY 297
Db 241 LLCFLPYHALRTLHLV---TWKDCGVDLHKATVITLTMAAANSCFNPFLYYFAGENF 296
QY 298 REMLISKF 305
Db 297 KARLRAIF 304
```

## RESULT 15

```
Q95LH1 PRELIMINARY; PRT; 355 AA.
AC Q95LH1;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE G protein-coupled receptor 1
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Wade-Evans A.M., Javan C., Russell J., Jenkins A.;
RT "Cloning and sequencing of simian G-protein coupled receptors, which
RT may function as SIV/SHIV co-receptors, from cynomolgus macaque
RT PBMCs."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292382; AAK97052.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 355 AA; 41370 MW; 1DDE41BE9CAC525C CRC64;
```

```
Query Match 19.2%; Score 316.5; DB 6; Length 355;
Best Local Similarity 29.2%; Pred.No. 2.6e-20;
Matches 86; Conservative 59; Mismatches 119; Indels 31; Gaps 9;
QY 26 YAIETIFGLGNTVVFVGLFCMKNNSS--SNVYLFNLSISDFAPLCTLPILIKSYA--N 81
Db 45 YCLSVLGIPIGNAIVW---FTGFKWKRTVSTLWFLNLAIADFIPLFLPLYI-SYVVMN 100
QY 82 DKGTGVDVLCISNRYVLTHTNLTSLFTFTFISMDRYLLMKYFPREHFLOKKEFAILISLA 141
Db 101 FHWPGIWLCKANSTAOLNMFASVFFLTVISLDHYTHLIHPVLVSHRHTLKNLSLIVIF 160
QY 142 VWALVTLVPLMFTFINSVPKKEGNSCIDYASSGNPEHNLISLCLT----LLGFLIPLS 197
Db 161 IWLLASLIGGPAFYFDVTEVFNHNTLCYNQFKDHPDLTVIRHVLVWVYVIGFLPPL 220
QY 198 VMCFYKMMVFLKRRSQOQATALPLDKPQRLVAVVI-FSILFTPYHIMRNRIASRL 256
Db 221 TMSICVLCILKLVKKRS-----ILISSRHFWTILAVVAVVVCWTPYHLSIWEITIH 274
```

```
QY 257 DSMPOGCTQKAISYITLTPRLAFLNSAINPIFYFLMGDHYREMLISKEROYFKS 311
Db 275 NSYSHHVMOAGI-----PLSTGLAFLNSCLNPILYVLIS-----KKFOARFRS 317
```

Search completed: April 2, 2003, 15:19:21  
Job time : 44 secs

